

PEPTIDES AND PROTEINS FOR EARLY LIVER DEVELOPMENT AND
ANTIBODIES THERETO

CROSS REFERENCE TO RELATED APPLICATIONS

This application is a divisional application of U.S. patent application serial no. 09/431,184, filed November 1, 1999, which is a continuation-in-part of PCT application PCT/US98/08656 which is a continuation-in-part of U.S. patent application Serial No. 08/841,349, filed April 30, 1997.

FIELD OF THE INVENTION

This invention relates to peptides and proteins isolated during early liver development, genes coding for these peptides and proteins, and antibodies which recognize these proteins, and to methods for their use in diagnosing and treating liver disease and other disorders.

BACKGROUND OF THE INVENTION

In the United States and other countries, end stage liver disease due to infection, genetic defects or alcoholic consumption is a major cause of widespread morbidity and mortality, causing great potential hardship and economic loss to millions of people throughout the world. In addition, numerous other diseases, including biliary problems and blood disorders, are associated with disruptions in the many functions carried out by the liver, including iron transport, hepatocyte formation and hematopoiesis. In general, severe problems associated with a breakdown of liver function are practically untreatable, and require a liver transplant as the only cure. However, in light of the great disparity between the number of patients needing liver transplants and the number of donors, thousands upon thousands of people are denied this operation, and transplantation is at the present time not a practical approach to the problem.

At the same time, the precise nature of liver development and the role of early developing liver proteins has not been well understood. To date, no growth factors specific to the liver have been identified or isolated, and the precise molecular mechanisms behind hepatocyte (liver cell) formation

to identify and understand the changes in gene regulation and expression in the developing liver, including the determination as to which genes are switched on and off as a hepatocyte forms and a liver develops. Accordingly, isolating and identifying the genes and proteins which play critical roles in early liver development would be beneficial in understanding the effect of gene regulation and expression in the differentiating liver, and in diagnosing and treating many diseases states involving the liver and liver functions.

SUMMARY OF THE INVENTION

Accordingly, it is thus an object of the present invention to provide genes comprising the nucleic acid sequences encoding early liver developmental proteins, including the liver proteins known as elf 1-3, liyor-1 (145), pk, protein 106, and praja-1.

It is further an object to provide isolated and purified early developing liver proteins encoding by the above genes.

It is still further an object to provide proteins which are characteristic of early liver development and peptides from said proteins and peptides, and to raise antibodies from said proteins and peptides which will be useful as markers, and will be useful in methods of identifying such peptides and proteins, tracing hepatocyte lineage, and treating liver disease.

It is still further an object to use the early developing liver proteins of the present invention to provide liver-specific growth factors for application in diagnosis and treatment of liver disorders.

It is still further an object to provide methods of diagnosing and treating end stage liver disease using the early developing liver proteins of the present invention.

It is even further an object to provide methods of diagnosing and treating other liver disorders and other diseases, including carcinoma, degenerative neurological disorders, anemia, and ataxia, using the early developing liver proteins of the present invention.

These and other objects are achieved by virtue of the present invention which provides genes coding for various proteins which are involved in the differentiation of the developing fetal liver, including the proteins known as elf 1-3, liyor-1 (145), pk, protein 106, praja-1, and a number of other stage-specific genes coding for early-developing liver proteins, and methods for their use in diagnosis and treatment of a variety of liver diseases and other disorders.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention will now be described in detail with respect to preferred embodiments thereof, which are to be taken together with the accompanying drawings, wherein:

FIGS. 1A-1B represent the nucleic acid sequence encoding the liyor-1 (145) protein in accordance with the present invention.

FIGS. 2A-2E represent the nucleic acid sequence encoding the elf-1 protein in accordance with the present invention.

FIGS. 2F-2I represent the nucleic acid sequence encoding the elf-2 protein in accordance with the present invention.

FIG. 2J represents the nucleic acid sequence encoding the elf-3 protein in accordance with the present invention.

FIGS. 3A-3B represent the nucleic acid sequence encoding the praja-1 protein in accordance with the present invention.

FIG. 4A-4B represent the nucleic acid sequence encoding the pk protein in accordance with the present invention.

FIG. 5 represents the nucleic acid sequence encoding the 106 protein in accordance with the present invention.

FIGS. 6A-6B represent the nucleic acid sequence encoding gene 20 in accordance with the present invention.

FIG. 7 represents the nucleic acid sequence encoding gene 36 in accordance with the present invention.

FIG. 8 represents the nucleic acid sequence encoding gene 41 in accordance with the present invention.

FIG. 9 represents the nucleic acid sequence encoding gene 112 in accordance with the present invention.

FIG. 10 represents the nucleic acid sequence encoding gene 114 in accordance with the present invention.

FIG. 11 represents the nucleic acid sequence encoding gene 118 in accordance with the present invention.

FIG. 12 represents the nucleic acid sequence encoding gene 129 in accordance with the present invention.

FIG. 13 is a depiction of the membrane skeleton of the ELF protein of the present invention.

FIG. 14 is a graphic representation of known alternatively spliced patterns found among *elf* transcripts.

FIG. 15 represents ELF expression in primary biliary cirrhosis.

FIG. 16 is a schematic view of the role of SMAD proteins as intracellular mediators of TGF- β and activins.

FIG. 17 depicts α -feto protein labeling cells of hepatocytic lineage in wild type vs. *smad2*^{+/-}/*smad3*^{+/-}.

FIG. 18 depicts apoptosis in *smad2* and *smad3* mutants.

FIG. 19 depicts rescue of the hepatic phenotype by culturing in the presence of HGF.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

In accordance with the present invention, early developing liver proteins and the genes coding for them have been isolated and sequenced, and these genes and proteins can be utilized to diagnose and/or treat a wide variety of liver disorders and other ailments. In general, the present invention arose from the investigation of liver formation during embryogenesis when the liver and other organs are in transition from an undifferentiated state to a differentiated one. This setting captures the phases of liver formation beginning with ordinary sets of endodermal cells. In addition, the early steps in tissue differentiation are closely related to the process of oncogenesis and tissue repair, and thus the isolated early developing liver proteins obtained in accordance with the present invention should have implications for diagnosis and treatment of a range of diseases from end stage cirrhosis to

hepatocellular carcinoma and many other disease conditions.

In the identification and isolation of the liver proteins of the present invention which are useful in early hepatocyte formation, the first step that was taken was to "capture" and analyze gene expression at different stages of early liver formation, particularly at those stages that emerge in the range of about days 9 through 14.5 in the mouse. In this regard, four embryonic liver cDNA libraries were constructed, such as at days 10.0, 11.5, 12.5 and 14.5 post coitus, and after subtractive hybridization, isolation of a group of stage-specific, liver restricted clones were isolated. As will be set forth in more detail below, sequence analysis has revealed that these clones encode a series of early developing liver proteins, which are generally "stage specific", i.e., they appear only at specific stages of development and not other stages, including elf proteins 1-3, liyor-1 (145), pk, protein 106, proteins coded for by genes 20, 36, 41, 112, 114, 118 and 129, and praja-1, as will be described further herein.

The initial project to identify and isolate developing liver proteins had four main objectives: (1) to construct early embryonic liver libraries; (2) to screen and characterize these early embryonic liver libraries with a group of probes comprising known growth factors (IGF-I, IGF-II, and IGFBP-2) and transcriptional activators (C/EBP and LFB1), known to be expressed in the developing liver; (3) to carry out subtractive hybridization utilizing these cDNA libraries and analyze subsequent subtracted clones for stage specificity by southern blot hybridization, sequence, transcript size, abundance, and tissue distribution; and (4) to develop a functional assay for these subtracted genes using embryonic liver explant cultures.

With regard to the main objectives of the invention, it was decided to focus on the four stages of liver development, particularly around days e10, e11, e12 and e14 (embryonic days post coitus) in developing mice. These four stages are defined developmental time points representing phases of liver development from undifferentiated mesodermal/endodermal cells

to a well developed and differentiated fetal liver. These stages have generally been categorized as follows: (1) at around e9-10, a change in cell polarity occurs; (2) at around e10.5-11, invasion and migration of endodermal cells into surrounding mesenchyme occurs; (3) at around e11.5-12, pseudolobule formation, cords of hepatocytes form together with early sinusoids; and (4) at around e12.5-e14.5, the liver is marked by hematopoietic foci and fully differentiated fetal hepatocytes. cDNA libraries representing these stages would therefore represent "captured" mRNA species expressed in greater abundance during critical time periods for hepatocyte formation, enabling their isolation and providing a method for analyzing the changing pattern of gene expression during liver development.

Another aspect of the present invention is the development of useful methods of diagnosis and treatment of liver disorders and other diseases made possible by the identification and isolation of the genes for early developing liver proteins of the invention and the expression of those genes. In accordance with the investigations made regarding these early developing liver proteins, it is clear that the different genes and proteins identified are important for different aspects of liver development and can thus be utilized in treatments of the appropriate disease. During embryogenesis, the liver generally develops from a foregut diverticulum, and comprises four main cell types: the first is the hepatocyte, or endodermal lineage; the second are biliary tree canalicular cells or bile duct cells, the third are hematopoietic cells, and the fourth are the Kupffer cell/Ito cells. As will be set forth below, of the early developing proteins isolated and obtained in accordance with the present invention, it is believed that the elf proteins are important for the formation of the biliary tree, as shown by antisense experiments; praja-1 appears to be important for iron transport and essential for hepatocyte formation as well as hematopoiesis; liyor-1 (145) and pk appear to be important in Ito cell formation and fibrosis.

Accordingly, in accordance with the invention, it is contemplated that elf proteins 1-3 will be useful in the treatment of disorders such as cholestasis, biliary stones, hepatic obstruction, stricture, primary biliary cirrhosis and primary sclerosing cholangitis. In addition, the proteins praja-1, liyor-1 (145) and pk will be useful in treating end stage liver disease, anhidrotic ectoderm dysplasia, hepatocellular carcinoma, as well as anemia, such as sideroblastic anemia, ataxia, e.g., spinocerebellar ataxia, degenerative neurological disorders, anhidrotic ectoderm dysplasia, and hemochromatosis.

Even further, it has also been discovered that the protein praja-1 has been identified in cancerous colon tissue, which normally does not produce this protein. Accordingly, it is contemplated that in accordance with the present invention, a method of detecting and diagnosing colon cancer is provided wherein colon cells or tissues are taken from a patient being tested, and these cells or tissues are screened to determine the presence or absence of the praja-1 protein. Identification of praja-1 in the colon cells or tissues will allow for a determination of whether the cells are cancerous since praja-1 will generally not be detectable in non-cancerous colon cells.

In the preparation of cDNA libraries in conjunction with this invention, it was necessary to utilize the four developmental stages discussed above in order to isolate key early developing liver proteins that affect the formation of hepatocytes and the liver. Although these studies were performed on mice, the relevance of the stages of liver formation to human development is shown in the following summary of these investigations:

(1) Day 10 post coitus (e10, 34-39 somites) (Human day 27):

In the mouse, the primary liver diverticulum appears during the 10th day of gestation. It develops from a foregut indentation in the endoderm which arises at e7, at the boundary between the embryonic and extraembryonic region, anterior to the developing heart rudiment. At this stage, although the

cells are committed to the formation of fetal hepatocytes, they are still epithelial in nature and the liver diverticulum is not viable in the absence of the surrounding heart mesenchyme.

As this is the earliest stage possible when hepatocytes are undifferentiated, it was considered to be of great importance: some cells are poised to differentiate into hematopoietic cells and others into hepatocytes. Accordingly, a day 10.0 library was constructed in lambda Unizap, and no prior group had ever constructed a cDNA liver library at this stage.

(2) Day 11.5 post coitus (e11.5, 40-44 somites) (Human day 32):

This stage is characterized by rapid growth of the liver.

Soon after the formation of the hepatic bud, the endodermal cells proliferate, disrupting the membrane separating the epithelium from the septum transversum, with the epithelial cells migrating into the mesenchyme. The liver at e11.5 consists of broad hepatic cords separated by large sinusoids containing nucleated erythrocytes. Hematopoietic foci are found intermingled with the hepatic cords. A cDNA library was constructed in lambda gt10 and lambda Zap from embryonic livers obtained at this stage, since although cells are proliferating rapidly, they still have not attained a fully differentiated fetal state.

(3) Days 12.5-13.0 post coitus (e12.5, Human days 35-45) (Embryo size: 7-9mm):

This stage is easily recognized by early signs of finger development as well as by the anterior indentation of the footplate. At this stage, the liver is well developed, all lobes being clearly visible; it contains many megakaryocytes as well as cells with erythropoietic activity. A cDNA library at e12.5 was constructed in lambda gt10 and lambda Zap as this was the earliest stage where fully differentiated fetal hepatocytes are seen.

(4) Day 14.5 post coitus (e14.5, Human days 51-57) (Embryo size: 20-32mm):

At this stage, individual, separated forefoot fingers can

be seen; hair follicles in the skin can be recognized and the umbilical hernia is very conspicuous. This stage represents a well differentiated fetal liver containing scattered blood-forming foci. A cDNA library of this stage was constructed in lambda Unizap in order to facilitate subtraction with the day 10 library which was also constructed in lambda Unizap (Stratagene).

(5) Adult mouse liver:

At birth, day 19, there is a major "switch" in the expression of a large number of genes. From now until the stage at which adult liver is formed, enzyme synthesis of the urea cycle and gluconeogenesis are upregulated. Adult liver is no longer a center for hemopoietic activity except in pathological situations and hepatocytes do not enter de-differentiated states, though the liver still has regenerative capacity as seen in partial hepatectomy.

In conjunction with each of these stages of development, RNA was recovered from each stage, and the quality of the RNA obtained following dissection was assessed by Northern blot analysis using mouse Beta Actin from the Chiba Cancer Center Research Institute, Chiba, Japan. Table 1 shows the RNA yields obtained. The cDNA library construction at days e11.5 and e12.5 of the embryonic liver was carried out by conventional techniques, and the libraries of the day e10.0 and adult mouse liver were obtained using the Stratagene Unizap cDNA library kit. The cDNA inserts for e10.5, e11.5, e12.5, e14.5 post coitus stage mice and the adult mouse liver were size selected on a Biogel A150 column (>500 bp) prior to ligation to the vector.

Qualitative analysis of cDNA libraries utilizing known developmentally regulated cDNAs were carried out in order to establish developmental profiles of important "early" genes that are significant in development, and these libraries were then screened with a specific number of probes. The following probes were obtained and used for screening these libraries, including: Insulin like growth factor I (IGF I), obtained from

Dr. Derek le Roith (NIH); Insulin like growth factor II (IGF II) and IGF II binding protein -2(BP- 2) both obtained from Dr. Matt Rechler of NIH; LFB 1 obtained from Drs. Monaci, Nicosia and Cortese of EMBL in Heidelberg, and the C/EBP probe from Dr. Darnell of the Rockefeller University in New York, N.Y.

The data shows that IGF I was not detected in any of the embryonic libraries, while IGF II was detected in increasing clone frequency from e6.5 to 8.5 (8 at e6.5, 8 at e7.5 and 38 at e8.5 - data not shown) and was also detected in the e10.0 and e12.5 libraries (3 at e10.0 and 4 at e12.5 - see Table 3). IGF II was not detected in the adult liver library. Interestingly, BP2 clone frequencies are similar to IGF II in the early e6.5, e7.5 and e8.5 libraries (data not shown), but in the liver cDNA libraries the clone frequencies differed, for BP2 only one clone per 100,000 being detected at e10.0 and e11.5, while 7 were detected in the Adult Liver Library compared to the greater numbers for IGF II. This implied that its temporal and spatial expression in the embryo and fetus is different from IGF II and this was subsequently confirmed by in situ studies. LFB I was detected in the e12.5 library, but at one clone per 100,000 screened, which implied that it is not present in mitogenic cells, but that its level was regulated and increased from birth onwards. C/EBP was not present in the e6.5, e7.5, e8.5 or e10.0 libraries (data not shown) but was suddenly detected at day e11.5 and e12.5 in low abundance (about 2 clones/100,000 at e11.5 and 5 at e12.5), implying that while it is expressed, its level also may be regulated, albeit downward, in embryonic stages. Lastly, Beta Actin was used as a reference: all seven libraries had similar Beta Actin frequencies from 120-300/100,000 clones which is considered representative of such embryonic libraries.

Next, identification of stage specific clones by subtractive methods was carried out, and two subtracted libraries were then constructed. Two rounds of subtraction were carried out, and the resulting subtracted libraries comprised 64 clones (e11.5-12.5), and 174 clones (e10.5-14.5).

Further characterization of these clones was carried out as follows: (1) Southern hybridization; (2) sequencing; (3) Northern analysis; (4) Zoo blot analysis; and (5) In vitro translation of protein.

In the Southern blotting of these clones, thirty-four clones were shown to be stage specific and not containing mitochondrial, ribosomal and globin sequences. DNA sequencing of these thirty-four stage specific clones was carried out in order to identify clones bearing homology to known developmental genes (such as cell polarity genes, homeobox genes, etc.), and the first 400 base pairs of each clone were sequenced. A detailed analysis was then carried out with respect to some of the clones which form a part of the present invention, including liyor-1 (145), protein 106, pk, and praja-1, since these clones exhibit true stage specificity and appear to belong to a set of genes encoding signal transduction proteins, which are of great interest in development currently, due to studies demonstrating their importance in cell lineage.

Other stage-specific proteins which are coded for by genes in accordance with the invention are discussed further below. Studies carried out with regard to proteins such as praja-1 and elf, as well as other early developing liver proteins, have elucidated the sequence of these proteins, as will be set forth in more detail below.

As an example of the tests used to elucidate the developmental expression of these liver proteins, the protein liyor-1 (145) was tested to determine whether these transcripts are differentially expressed during development, specific for mesoderm or endoderm derived tissues, or are expressed in adult mouse and human organs. Accordingly, tissues from mid-gestational embryos were analyzed to determine the role of 145 in liver development. In these tests, tissues were dissected from day 11 onwards, as it was at this stage that discrete hepatic, cardiac and other tissues could be dissected with ease, with subsequent RNA isolated being of good quality.

RNA hybridization with liyor-1 (145) DNA in different

mouse tissues was studied by using polyA RNA obtained at various developmental stages using a ^{32}P -labeled 1.1 Kb insert representing protein 145. The specificity of the developmental changes in the steady state levels of 145 was evaluated by also measuring the relative levels of Actin. This revealed a 2.4 Kb transcript at high stringency washes. Scanning densitometry of the respective bands revealed that maximal expression of 145 occurred in liver and heart, less so in other tissues, but specifically on day 11 and in decreasing abundance at days 12.5 and 14.5 (when Northern blots were developed 1-2 months later).

Further characterization of the distribution of protein 145 RNA in adult tissues and its conservation in evolution has involved RNA analysis of adult mouse and human tissues. The protein 145 hybridizes to adult liver, kidney and testis as a 2.4 Kb transcript in liver and kidney and a 2.6 Kb transcript in adult testis, in very low abundance: both blots were developed after being exposed to film for over a month at -70°C .

Similar tests conducted with regard to the elf protein and the nucleic acid coding for it showed that elf DNA is generally conserved across many different species, including human, monkey, rat, mouse, dog, cow, chicken and yeast, and is represented in all species studied except rabbit.

Finally, in accordance with the invention, a functional assay was established for subtracted genes with the goal to establish mouse embryonic liver explant cultures in the laboratory, as this is usually considered the major hurdle for antisense experiments due to the need to dissect extremely small tissue sections at day 9.5 when the liver bud is 0.2 mm.

In this regard, the interactions of the neighboring cardiac mesoderm and foregut endoderm were studied and the subsequent changes in cell type specific gene expression were characterized, particularly with respect to alpha-fetoprotein and albumin expression, and partially with respect to epithelial basement membrane components. Methods of culturing liver explants in accordance with the invention are described below. The results obtained in these tests have shown that

when cultured in the complete absence of mesodermal derivatives, hepatic endoderm deteriorates rapidly. Only 2 out of 15 such liver explants survived. Hematoxylin and eosin staining showed a necrotic endoderm with no apparent signs of hepatic differentiation. When associated with the surrounding mesoderm particularly cardiac mesoderm (en bloc dissections), the endodermal cells had proliferated and invaded the mesodermal strands. Hepatocytes were seen to be organized in cords separated by sinusoids with pseudo-lobule formation. All 15 out of 15 cultures from en bloc dissections were completely viable. These studies confirm prior explant studies demonstrating the necessity of surrounding mesoderm for liver formation.

Accordingly, cDNA libraries have been constructed for the four main stages of liver development, e10, e11.5, e12.5, e14.5 and for adult liver in the mouse. These have been shown to be truly representative of their respective MRNA species, by meticulous analysis utilizing initial RNA blot analysis, size fractionation, quantitative, and qualitative analysis. Northern analysis confirmed the stage specificity, and restricted expression of their transcripts: for 145 this comprised a 1.35, and 2.37 Kb transcript restricted to midgestational brain and liver tissue, and adult mouse and human Northern blot analysis revealed 145 transcripts in extremely low abundance in liver, kidney, testis. Further tests with regard to protein 145 reveals its sequence identity of 53% (20 S.D.'s) to rat Phospholipase C- γ (PLC- γ), and amino acid alignment of conserved section of 145 to PLC- γ identifies a split pleckstrin homology (PH) domain. Protein 145 (liyor-1) bears 99% identity at the amino acid level to the PH domain at the amino terminus of PLC- γ . The PH domain is an area of 100 amino acids that has been found in a number of proteins including serine/threonine kinases, GTPase activating proteins, phospholipases and cytoskeletal proteins, and is thought to be involved in signal transduction. Nuclear magnetic

resonance spectroscopy has revealed that the PH domain of P-fodrin is an electrostatically polarized molecule containing a pocket which may be involved in binding of a ligand. Of immense interest is the fact that this pocket is related to the peptidyl-prolyl-cis- trans-isomerase FKBP in which this pocket is involved in the binding of the macrocyclic compound FK506. Accordingly, it is contemplated that protein 145 may indeed bear a pocket for 'natural' ligand similar to FK506 and thus appears to be a potential factor for hepatocyte differentiation.

PLC- γ is regulated by a combination of SH2- domain dependent complex formation with tyrosine phosphorylated receptor tyrosine kinases, and its subsequent phosphorylation on tyrosine residues. An unique feature of PLC- γ and protein 145 is that both contain a split PH domain, which in the case of the PLC fills the gaps between the SH2-SH2-SH3 region and the surrounding X and Y catalytic domains. The SH2-domains mediate the high affinity interaction of PLC- γ with activated growth factor receptors such as epidermal growth factor (EGF) or platelet derived growth factor (PDGF) receptor. The PH domain similarly may be utilized as a specialized noncatalytic domain directing complex formation between protein kinases and their presumptive targets during liver development. In addition, the area of complete identity and split PH domain in 145 and PLC- γ is conserved in a number of other proteins through to TOR2, an essential yeast PI 3 kinase, and to v-abl. A parallel can be drawn to the SH2 domain: that proteins associating with activated growth factor receptors have quite distinct enzymatic properties, are structurally unrelated within their catalytic domains, yet contain a similar noncatalytic domain of approx 100 amino acids, called the src homology (SH) region 2. The SH2 domain was first identified in non receptor protein tyrosine kinase like Src and Fps, by its apparent ability to interact with the kinase domain and phosphorylated substrates. It is believed that during the

evolution of cellular signaling mechanisms, the acquisition of SH2 domains conferred on PLC- γ and GAP the capacity to interact with transmembrane tyrosine kinases and therefore couple growth factor stimulation to PI turnover and the kinase pathway. PH domains are similarly conserved and may be utilized in the same way that SH2 domains are.

As indicated above, the protein liyor-1 (145) appears to be important in Ito cell formation and fibrosis, and is thus thought to be useful in treating end stage liver disease as well as other conditions including hepatocellular carcinoma, anemia, ataxia, and hemochromatosis. It is contemplated that the use of the protein Liyor-1 will be by administering to a suitable patient an amount of this liver protein effective to treat the specific condition of that patient, and this would be carried out using conventional means and regimens well known to one skilled in this art. The sequence of Liyor-1 which was determined using the cDNA libraries of the present invention is shown in Figs. 1A-1B, and suitable amounts of the liyor-1 (145) protein may be prepared in a conventional manner by expressing by recombinant or other means the nucleic acid coding for the 145 protein, after which the protein can be isolated and/or prepared into substantially pure form as needed. In addition, the 145 protein may be administered with any other suitable compound normally utilized for administration into a patient, such as a suitable pharmaceutically acceptable carrier.

As indicated hereinbelow in the examples, other genes for early developing liver proteins in accordance with the present invention have been isolated and sequenced, including the genes coding for the elf proteins, praja-1, pk protein, protein 106, and genes 20, 36, 41, 112, 114, 118 and 129. With regard to the elf proteins, these proteins were studied by analyzing mRNA from tissues from mid-gestation embryos. Tissues were dissected from day 11 onwards, as it was at this stage that discrete hepatic, cardiac and other tissues could be dissected with ease, and the subsequent RNA that was isolated was of good quality. RNA hybridization with elf DNA in different mouse

tissues was studied by using polyA RNA obtained at various developmental stages using a ^{32}P -labeled 1.1 Kb insert representing elf. The specificity of the developmental changes in the steady state levels of elf was evaluated by also measuring the relative levels of Actin. This revealed a 2.4 Kb transcript at high stringency washes. Scanning densitometry of the respective bands revealed that maximal expression of elf occurred in liver and heart, less so in other tissues, but specifically on day 11, and in e12.5 and e14.5 in decreasing abundance (when Northern blots were developed 1-2 months later).

In situ hybridization was then used to confirm elf expression in 11.5 heart and liver as well as to determine its expression pattern during earlier liver development, as will be set forth below in the Examples. The liver bud, which originates from foregut endodermal cells, grows into the septum transversum at the 9th day of gestation (13-20 somite stage). Between days 10.5 to 11.0 post coitus, a considerable degree of differentiation occurs: The liver enlarges substantially over this period, this increase in volume being due to the invasion of the mesenchyme of the septum transversum by the hepatic cords, and the initiation of hematopoietic activity in the liver. At day 9.5, a strong labeling of elf becomes apparent in the heart, and the pattern appears to be trabecular, including the wall of the cardiac anlage. A section of the sino-atrial chamber wall also shows a high intensity of elf expression. The surrounding tissue, particularly the caudal liver bud region does not show the presence of silver grains.

At the next stage, day 10.5, silver grains clearly highlight the developing liver, which appears as a horizontal structure (L) in this section. At this stage, the signalling is weakening in the developing heart tissue. The surrounding tissues are remarkable for the absence of silver grains. At day 11.5, a strong labeling of elf becomes apparent in the liver, which is larger in size. The heart at this stage only shows a weak signal posteriorly. As a control, in addition to sense probes, a riboprobe to alpha fetoprotein outlines the

developing embryonic liver at days 11-12.

A comparison of the day 9.5 and 10.5 embryos, demonstrates a temporal and spatial expression of elf: the temporal gradient of a rise and fall of elf expression in the heart can be inferred from the strong staining in the developing heart at day 9.5 followed by a weaker staining at the next stage (day 10.5). Simultaneously, liver expression increases. The spatial gradient is apparent where silver grains increase in density on moving from the developing heart to the liver: at day 10.5, antisense RNA probes from elf cDNA hybridized specifically to day 9.5 cardiac mesenchymal tissue; expression at day 10.5 being restricted to cardiac and hepatic tissue, with elf expression finally being restricted to the liver in later 11.5 day embryos. Of note, elf expression was seen in embryonic livers at later stages (days 12.5, 14.5 p.c.), but only in decreasing abundance: the message being detected in these later stages when Northern and in situ were developed a considerable time later. Sense probes to elf did not hybridize to any tissues. This indicates that ELF expression is not a sudden "on" "off" phenomenon, but more of a gradient pattern: consistent with the expression pattern of brain beta spectrin.

Alpha fetoprotein antisense RNA probes hybridized specifically to 11.5, 12.5, 14.5 embryonic mouse liver tissue, which is in agreement with previous studies of mRNA isolated from embryonic liver samples. The earliest stage that we were able to detect alpha-fetoprotein mRNA by in situ hybridization was at 10.5-11.0 days of gestation. Similar experiments with albumin mRNA have shown it to be expressed at day 9.5 in clusters of cells arising from foregut epithelium and in cords of cells seen to be invading the septum transversum. In experiments with alpha-fetoprotein, the liver was labeled at all subsequent stages (day 11 onwards), and, upon histological examination appeared to occur primarily in the endothelial cells. Hematopoietic cells appeared retractile but did not contain the hybridization grains that were visible over the alpha-fetoprotein positive cells. These experiments show that

elf mRNA is localized to early embryonic heart, and then moving to ell liver.

Next, it was determined that elf was a marker for the mesodermal component of liver formation. As Northern analysis had revealed elf expression to occur in day 11.5 heart and liver tissue, in situ localization was performed to investigate whether elf expression was restricted specifically to mesodermal tissue from the heart and the liver and was then compared to the endothelial expression of alpha fetoprotein. The main regions of mesoderm in the developing embryo are dorsal (somitic), intermediate, and lateral. Specifically, lateral plate mesoderm comprises somatic tissues (pleura, pericardium, peritoneum and limb bud), and splanchnic tissues (heart, epicardium, myocardium, connective tissue and smooth muscles of viscera and blood vessels, hemangioblastic tissue, adrenal cortex and spleen). The developing heart, at day 9 (13-20 somites), appears to be only region within the embryo where the endothelial elements of the circulation are surrounded by a vessel wall. The walls of the common ventricular and atrial chambers show an increasing degree of trabeculation. The space between the endothelial and myocardial elements is filled with loose mesenchyme called cardiac jelly. In situ hybridization of days 9 and 10 embryonic heart tissue using elf antisense riboprobes showed high levels of labeling to both the atrial and ventricular regions, highlighting the trabeculation.

Hepatic mesenchyme also originates from lateral plate mesoderm. The septum transversum part of the hepatic mesenchyme originates from the splanchnic mesoderm of the precardiac area and this is considered to be responsible for the subsequent differentiation of hepatocytes. However, tissue explant experiments have demonstrated that all derivatives of the lateral plate can replace hepatic mesenchyme for these later events. The initial experiments have shown that migrating endoderm must interact with mesenchyme for the former to differentiate into hepatocytes and recent studies

investigating albumin mRNA expression, an indicator of hepatocyte differentiation, have confirmed these features; Initial expression of albumin mRNA occurs during the invasion of the septum transversum, when foregut endodermal cells clearly contact cardiac mesenchymal tissue. Similarly, primer extension analysis of albumin transcription has shown that the start site of transcription to occur at day 10.5 with a 15-20 fold increase in albumin mRNA upon liver organ formation by day 12.5. In our experiments using alpha fetoprotein as a marker for differentiated hepatocytes, it was obvious that while alpha fetoprotein expression is restricted to the later endodermal component of liver development, elf expression seems to occur in the loosely organized, lighter staining mesenchymal cells, initially cardiac mesenchyme (at day 9.5), then in both cardiac and hepatic tissue (at day 10.5) and then restricted to liver tissue (day 11.5 onwards); elf expression then decreases in abundance upon full embryonic liver formation. Examination of later histological sections (days 11 onwards) demonstrated a diffuse distribution of grains, and the hybridization signal with elf appeared to be localized in the perisinusoidal cells, but not in the hepatocytes.

That elf is expressed in early cardiac mesoderm, with subsequent expression being limited to hepatic mesoderm, indicates that this is a novel marker for the mesodermal component of liver development. Molecular markers have been invaluable in the dissection of inductive events in embryological studies. For instance, in Xenopus, vg-1, a member of the TGF-Beta family, now considered to be the strongest candidate for dorsal mesoderm induction, was in fact originally isolated by differential screening of mRNAs localized in the vegetal hemisphere of developing Xenopus eggs.

Activins and other genes belonging to the TGF-Beta family such as vg-1, as well as wnt and BFGF families, represent components of the cascade leading to the commitment to particular mesodermal fate and all are strong candidates as mesoderm-inducing factors. Yet of these, only vg-1 has been

demonstrated to be localized to the vegetal cells, the blastomeres responsible for mesoderm induction in vivo. Specific localization of vg-1 was vital and responsible for the persistence required in investigating its role as the inductive agent in mesoderm formation. Similarly, in isolating putative inductive agents required for liver formation, a key step is the localization of a new mRNA isolated from the embryonic livers. Accordingly, it is contemplated that elf and its associated regulatory genes will be of enormous potential benefit as a liver growth factor.

Further characterization of elf has involved RNA analysis of adult mouse and human tissues, and it was determined that elf hybridizes to adult liver, kidney and testis as a 2.4 Kb transcript in liver and kidney and a 2.6 Kb transcript in adult testis, in very low abundance: both blots were developed after being exposed to film for over a month at -70°C . Genomic DNA analysis of elf expression in DNA (genomic) from human, monkey, rat, mouse, dog, cow, rabbit, chicken and yeast indicates that elf is conserved across the species, being represented in all except rabbit DNA.

In vitro transcription and translation of elf, the latter using nuclease-treated rabbit reticulocyte lysate (promega) has revealed a 34 Kd protein, which is as predicted by the elf insert size and indicating that this insert is in frame for the coding sequence for a specific protein. These studies have established the principle that specific mesodermal mRNAs are localized in a way that guarantees their subsequent segregation to specific mesodermal tissue, in this case the presumed mesodermal component of the liver as shown by embryonic explant studies.

The elf protein has been sequenced, and it has been determined that at least three specific elf protein genes can be identified during early liver development. The sequences for these genes, known as elf-1, elf-2, and elf-3, are shown in the Figs. 2A-2E, 2F-2I and 2J, respectively. As indicated above, it appears that the elf proteins 1-3 are probably

important for the formation of the biliary tree during early liver development. Accordingly, it is contemplated that in accordance with the present invention, the elf proteins will be useful in treating various disorders associated with liver function, including cholestasis, biliary stones, obstruction, stricture, primary biliary cirrhosis, and primary sclerosing cholangitis. As would be readily apparent to one skilled in the art, methods of treatment using the elf proteins would comprise administration of an amount of an isolated elf protein that is effective to treat the specific disease condition described above. As also would be apparent, the elf proteins themselves can be prepared in a number of suitable ways by expression from the nucleic acid sequences indicated at Figures 2A-2J, including recombinant methods of producing these proteins, followed by separation, isolation and/or substantially purifying the elf proteins. The elf proteins once obtained in this manner can be put into any suitable form that is acceptable for use with patients. In addition, any of these three elf proteins may be administered with any other suitable compound normally utilized for administration into a patient, such as a suitable pharmaceutically acceptable carrier.

Another protein that has been identified and isolated in accordance with the present invention and which is contemplated to be used in a variety of therapeutic methods is known as praja-1. Praja-1 has now been studied in conjunction with the examination of early developing liver proteins, and an analysis of the amino acid translation revealed the presence of a COOH-terminal RING-H2 motif, which is a zinc finger variant. Additionally, Northern blot analysis of RNA from adult mouse showed expression of 3.1, 2.6, and 2.1 kb transcripts in liver, brain, and kidney, and an additional 2.3 kb transcript in testis. Expression of praja-I is also apparent in a colon cancer cell line, SW 480, and as set forth below, it is also contemplated that the praja-1 protein will be a useful marker in early detection of colon cancer.

It has also been learned that praja-1 maps to chromosome X, at about the 36 cM position. Other genes mapping to this general region include moesin (Msn), androgen receptor (Ar), interleukin-2 receptor gamma (IL-2rg), X-linked zinc finger protein (Zfx), and tabby (Ta). The syntony and conserved gene order between mouse and human X chromosomes allows comparison with human disease genes in the region. Human diseases in this region with mesodermal involvement include anhidrotic ectoderm dysplasia (eda) and sideroblastic anemia with spinocerebellar ataxia (asat), and it is thus contemplated that in accordance with the present invention, praja-1 will be useful in treating these disease conditions, as well as degenerative neurological disorders.

In vitro expression of praja-1 has shown that the translational product, which ran as two closely spaced bands of Mr= 55.6 and 56.9 kD, is larger than the predicted ORF size of 47.4 kD. One possible explanation is that the expression product is very acidic, and acidic proteins such as granins are known to give anomalously high Mr on SDS-PAGE. The presence of two products suggests translation initiation at a second, internal ATG codon, such as at Met-19.

In addition, antisense studies to praja-I demonstrated that praja-I is essential for liver architecture formation. Preliminary antisense studies were performed at 1.25, 2.5 and 5 mfn concentrations, utilizing two different ODNs to praja-1. In these tests, liver and block explants were treated with these antisense ODNs compared with control (scrambled, sense or no ODNs). The results showed that control livers were generally larger than the antisense-treated livers, and control blocks showed early hepatocyte growth, cartilage growth, and very preserved bile ducts. Both livers and blocks treated with either antisense ODN to praja-1, showed minimal hepatocyte growth, cell necrosis, yet preservation of cartilaginous tissue, in a dose dependent manner.

In praja-1, aside from the RING-H2 finger, the stretch of thirty-four COOH-terminal amino acids just past this motif is

especially rich in proline residues (17.6%); and, as stated, the protein in general is very acidic. Proline-rich domains are found in several mammalian transcription factors, such as that at the COOH-terminus of transcription factor CTF. Proline-rich regions and also acidic regions are likely to function in contacting other proteins. When considering the praja-I sequence as a whole, the rat Neurodap1 gene has the highest similarity. Neurodap1 is expressed abundantly in rat brain, with much smaller amounts in heart and skeletal muscle.

Though praja-I likewise shows expression in brain, unlike Neurodap1 (which is a larger 4.8 Kb transcript), it also expresses in liver and kidney. The subcellular localization of Neurodap1 was shown to be concentrated around the endoplasmic reticulum (ER) and golgi of the cerebral cortex and facial nucleus, and especially in the postsynaptic density region of axosomatic synapses. Based on its subcellular localization, plus the presence of the RING-H2 finger, Neurodap I is probably linked to the secretory or protein sorting. This similarity to Neurodap1 indicates that praja-I is most likely involved in protein-protein interactions, possibly in a protein sorting or secretory pathway involved during hepatocyte formation.

The gene coding the praja-1 protein has been sequenced, and this nucleic acid sequence is shown in Figs. 3A-3B. As indicated above, it appears that the praja-1 protein is probably important for iron transport, and essential for hepatocyte formation as well as hematopoiesis. Accordingly, in accordance with the present invention, praja-1 can be used in methods of diagnosing and treating diseases such as end stage liver disease, iron storage disorders, hepatocellular carcinoma, as well as anemia, such as sideroblastic anemia, ataxia, such as spinocerebellar ataxia, and hemochromatosis. As would be recognized by one skilled in the art, these methods of treatment would involve administering of an effective amount of the praja-1 protein to the patient afflicted with one of the disease conditions set forth above. In addition, the isolation of the praja-1 protein could be obtained by expression of the

nucleic acid sequence indicated at Figs. 3A-3B which codes for the praja-1 protein, and this protein can be produced from its nucleic acid sequence in any suitable manner well known in the art such as recombinant means. Once isolated in this manner, the praja-1 protein can be obtained in a desired form, such as in substantially purified condition, and can be incorporated into any suitable mode of treatment that would be compatible with the patient in need of such treatment. In addition, the praja-1 protein may be administered with any other suitable compound normally utilized for administration into a patient, such as a suitable pharmaceutically acceptable carrier.

Even further, as indicated above, it has also been discovered that the protein praja-1 has been identified in cancerous colon tissue, such as in colon cancer cell line SW 480, which normally does not produce this protein. Accordingly, it is contemplated that in accordance with the present invention, a method of detecting and diagnosing colon cancer is provided wherein colon cells or tissues are taken from a patient being tested, and these cells or tissues are screened in any suitable manner which would identify the presence or absence of the praja-1 protein in the tested cells or tissues. In this manner, the identification of praja-1 in the colon cells or tissues from the patient will be indicative of a cancerous condition in the colon cells or tissues, and thus the present invention will provide a simple and effective method for determining at an early stage, when the disease is still in a treatable condition, if the patient appears to have contracted colon cancer. Conversely, the absence of praja-1 will generally be indicative of a non-cancerous state in tested colon cells.

Still other genes coding for early developing liver proteins in accordance with the present invention have been identified and sequenced, and these proteins will also be useful in various methods of diagnosis and treatment of disease conditions associated with the liver or liver function. Included in these additional genes are those nucleic acids

coding for a protein identified as pk, as depicted in Figs. 4A-4B, nucleic acids coding for a protein identified as protein 106, as shown in Figure 5, and genes 20, 36, 41, 112, 114, 118 and 129, as shown in Figures 6-12. These proteins also appear to useful in hepatocyte formation and in treating liver diseases in a similar manner to many of the proteins discussed above, and in a manner similar to known growth factors should be useful in treating a variety of conditions. For example, protein pk appears to be important in Ito cell formation and fibrosis and thus appears to be useful in the same manner as protein liyor-1 (145). Accordingly, the protein pk, as prepared from the nucleic acid sequence indicated at Figure 4, will likely be useful in treating end-stage liver disease, hepatocellular carcinoma, as well as other disease conditions including anemia, ataxia, and hemochromatosis. As in the above cases, these early developing liver proteins may be administered with any other suitable compound normally used for administration to patients, such as suitable pharmaceutically acceptable carriers.

Another aspect of the present invention will comprise raising antibodies to the early developing proteins identified above, or to peptides or fusion proteins such as the pk protein (also known as itih-4) derived from these proteins. As would be recognized by one of ordinary skill in the art, antibodies to these proteins or to selected peptides or fusion proteins derived from these proteins may be prepared in any suitable conventional manner currently known, including raising antibodies in such animals as rabbits, sheep, goats, or guinea pigs. In the preferred embodiment, the following antibodies have been raised in rabbits:

- (1) Peptides (13-mer) at aa 2-14 of mouse elf gene N-terminus having the sequence 5-ELQRTSSVSGPLS-3.
- (2) Peptides (14-mer) at aa 2140-2154 of mouse elf gene C-terminus having the sequence 5-FNSRRTASDHSWSG-
- (3) Peptides (13-mer) at aa 144-156 of mouse prajal gene middle portion having the sequence 5-

LRRKYRSREQPQS-3.

In addition, the invention also comprises antibodies to the following peptides:

(1) 145peptide-A (18-mer) which was designed from the C-terminus of gene 145 (Cded) and has the sequence 5-SAQSLVVTLGRVEGGIRV-3 OR 5-CSAQSLVVTLGRVEGGIRV-3.

(2) 145peptide-B (17-mer) which was designed from the middle part of gene 145 (Cded) and has the sequence 5-KIEGSSKCAPLRPASRL-3 or 5-CAPLRPASRLPASQTLG-3.

(3) g59peptide-A (16-mer) which was designed from the N-terminus of gene G59 (Praj1) and has the sequence 5-PPREYRASGSRRGMAY-3 or 5-PPREYRASGSRRGMAYC-3; and

(4) g59peptide-B (15-mer) which was designed from the middle part of gene 59 (Praj1) and which has the sequence 5-CKVPRRRRTMADPDFW-3.

The invention also comprises antibodies to a fusion protein such as a 40 kD pk/itih-4 fusion protein which covers the two EF-hands motifs of itih-4 (about 400-bp 14kD).

The invention further comprises the use of the elf proteins of the present invention in interactions with TGF- β signaling molecules such as Smad2 and Smad3 so as to prevent or treat liver diseases such as primary biliary cirrhosis (PBC) and other diseases involving bile ducts. Evidence has shown that SMAD2 and SMAD3 insufficiency leads to a loss of bile ducts, and thus that TGF β treatment of normal livers results in an increase in bile duct formation via the activation of SMAD2 and SMAD3. SMAD2/3 activity may be mediated by ELF, a Beta Spectrin. Loss of ELF function results in T lymphocytic proliferation and absent intrahepatic bile ducts. Livers deficient in SMAD2 and SMAD3 exhibit perturbations in ELF localization. This phenotype is seen in Primary Biliary Cirrhosis (PBC), a cholestatic disease with a progressive loss of intrahepatic bile ducts. Perturbations in ELF are

correlated with a lack of SMAD2 and SMAD3 in this disease. Immunoprecipitation studies show that ELF binds SMAD2 and SMAD3, and that this binding is increased in PBC.

Previously, it was observed that compound haplo-insufficiency at the *smad2* and *smad3* loci resulted in a failure to form intrahepatic bile ducts, and that HGF could rescue this phenotype in a SMAD-independent pathway. It was noted that TGF β could rescue the bile duct insufficiency in the mutant livers, although it did not completely rescue the hepatocytic defects seen. The effect of TGF β on the wild-type explants was also quite interesting. Treatment of fetal livers with exogenous TGF β in vitro resulted in a marked increase in the number of intrahepatic bile ducts. Moreover, the morphology of the bile ducts formed underwent a dramatic change. The bile ducts increased to twice their normal size, and were less regularly organized than those found in untreated liver explants.

In addition, it has also been previously shown that HGF was able to direct the formation of bile ducts while bypassing SMAD activation. The question of whether SMAD2 and SMAD3, the pathway specific SMAD proteins downstream of TGF β and activins, were activated in the livers in the presence of TGF β was also examined by looking at the subcellular localization of these SMAD proteins by immunofluorescence and confocal microscopy in explant livers from *smad2*^{+/-}; *smad3*^{+/-} mutants cultured in the presence of TGF Beta. It was determined that a narrow expression domain of SMAD2 is found just adjacent to the developing bile ducts, in which some SMAD2 appears to be nuclear, suggesting that it is activated and is transducing the TGF β signal. SMAD3 was also expressed adjacent to the forming bile ducts, although its expression domain was much wider than seen for SMAD2. SMAD3 could also be found in the nuclei of some cells, suggesting it too was activated in response to the TGF β . Therefore, diminution of SMAD activity through genetic haploinsufficiency ablates bile duct

formation, while exogenous activation of SMAD2 and SMAD3 can augment bile duct development, suggesting a central role of TGF β and SMAD2 and 3 in bile duct formation.

The phenotype of the *smad2*^{+/-}; *smad3*^{+/-} embryos was reminiscent of what was seen when ELF3 was inhibited by antisense oligonucleotides in liver explants (Mishra, Oncogene, 1999). Specifically, bile ducts failed to develop, and the hepatocytic architecture was highly deranged. ELF3 is a β -spectrin protein in accordance with the present invention, which we have shown is expressed in the membrane of hepatocytes. Indeed, recent confocal experiments have pinpointed a focused concentration of ELF3 protein at the apical, canalicular surface of the hepatocytes, which led to an examination of the effect of *smad2/3* ablation on ELF localization. The evidence regarding the association of *smad2* and 3 and the ELF proteins of the present invention is discussed in more detail in Example 3.

It is thus submitted that the foregoing embodiments are only illustrative of the claimed invention, and alternative embodiments well known or obvious to one skilled in the art not specifically set forth above also fall within the claimed scope.

In addition, the following examples are presented as illustrative of the claimed invention, and are not deemed to be limiting of the scope of the invention, as defined by the claims appended hereto, in any manner.

EXAMPLE 1

In accordance with the cloning strategy of the present invention to identify genes involved in early mouse liver development, we have isolated Praja-1, a gene with similar sequences to the *Drosophila melanogaster* gene *goliath* (gl), and which is involved in the fate of mesodermal cells ultimately forming gut musculatures, fat body, and the heart. Praja-1 is a 2.1 kb gene encoding a putative 423 amino acid ORF and includes a COOH-terminal RING-H2 domain. Using the Jackson Laboratory BSS panel, we have localized praja-1 on chromosome X at 36 cM, near the X inactivation center gene, Xist. Northern blot analysis demonstrated three transcripts (3.1, 2.6 and 2.1 kb) in mRNA from adult mouse tissues brain, liver, and kidney as well as in mRNA from developing mouse embryos (days 7, 11, 15 and 17 post coitus, or p.c.). In vitro transcription/translation yielded two products with a Mr of 55.6 and 56.9 kD. The presence of the RING-H2 domain, a proline-rich region at the COOH-end, and regions rich in acidic amino acids, leads to the hypothesis that the Praja-1 product is involved in mediating protein-protein interactions, possibly as part of a protein sorting or transport pathway. This is strengthened by the similarity of praja-1 to rat Neurodap1, whose product has been shown to localize to the endoplasmic reticulum and golgi in brain.

The molecular mechanisms underlying hepatocyte differentiation are not well understood, and thus identifying the genes underlying the control of liver development will provide powerful tools for understanding liver function and development, and will allow the use of inducing liver differentiation for therapeutic purposes. As part of a strategy to clone such genes, we isolated a new RING-H2 finger gene, praja-1. RING-H2 fingers, a type of zinc finger, are similar to RING fingers except that Cys4 is replaced by His (see Freemont, Ann. N.Y. Acad. Sci. 684:174-192 (1993); Lovering et al., P.N.A.S. 90:2112-2116 (1993)). Here we show that praja-1 possesses a RING-H2 motif near the COOH terminal.

The RING-H2 motif is similar to that of the *Drosophila melanogaster* gl gene (Bouchard et al., Gene 125:205-209, 1993), and to the rat Neurodapl gene (Nakayama et al., J. Neurosci. 15:5238-5248, 1995). Praja-1, which localizes to chromosome X, is expressed in mouse brain, liver, and kidney. The presence of the RING-H2 motif, plus the acidic, hydrophilic nature of the translation product, leads to the hypothesis that praja-1 plays a role in protein transport.

Materials and Methods

cDNA preparation and 3'-RACE PCR: RNA was isolated from livers of day 11 p.c. embryonic mice (ICR, Harlan Sprague-Dawley) using guanidine thiocyanate (Chomczynski et al., Ann. Biochem. 162:156-159, 1987). Poly(A)+ mRNA was isolated from total RNA using Dynabeads, as per manufacturer's instructions. First strand cDNA was made from poly(A)+ mRNA using the Promega Reverse Transcriptase System and the 3'-RACE primer 5'-GACTCGAGTCGACATCGA-T17 (Frohman, In: M. A. Innis et al. (eds.), PCR protocols: a guide to methods and applications, Academic Press, San Diego, pp. 28-38., 1990). The 3'-RACE primer was also used as the reverse primer in the PCR reaction.

The forward PCR primer, originally designed to amplify a conserved region of a clone 145/PH (pleckstrin homology) domain, was 5'-CTCAAGCAGGTCCTGGCACA. The PCR reaction mix contained cDNA from about 10 ng of poly(A)+ mRNA, 25 pmol of each primer, 1 mM dNTP mix, and 2.5 units of AmpliTaq DNA polymerase (Perkin-Elmer) all in 10 mM Tris, 1.5 mM MgCl₂, and 75 mM KCl, pH 9.2 in a final volume of 50 µl. The temperature program comprised 35 cycles of denaturation (94°C, 1 min), annealing (55°C, 1 min), and extension (72°C, 3 min), followed by an additional 8 minute extension. One of the resulting PCR products (CH7) comprised a 725 bp fragment, which was cloned into vector PCR11 using the Invitrogen TA Cloning Kit for sequencing, and found by sequence analysis to possess a RING-H2 finger. The portion of the final cDNA clones which correspond to CH7 is indicated in Figure 3.

Library screening: The PCR product CH7 was labeled with [α - 32 P]-dCTP (3000 Ci/mmol, Amersham) via primer extension using the reverse PCR primer plus AmpliTaq polymerase at 72°C in PCR buffer (Konat et al., in PCR Technology: Current Innovations (H.G. Griffin and A.M. Griffin, Eds.), CRC Press, Boca Raton, pp. 37-42, 1994). The resulting antisense probe was used to screen plaque lifts of a whole embryonic mouse (day 11 p.c.) cDNA library in vector λ Zap (Stratagene). Positive plaques were picked and purified, and DNA was isolated from lysates using standard procedures (Silhavy et al., Experiments with gene fusions, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1984). Inserts were excised from λ Zap DNA using EcoRI, and were subcloned into pGEM3Zf(-) (Promega) for sequencing and subsequent manipulations.

DNA sequence analysis: DNA sequence comparisons to existing sequences were performed utilizing BLAST searches in Genbank. Alignments were performed using the GCG program PILEUP.

Chromosomal mapping: Southern blot analysis of genomic DNA from C57BL/6J (B6) and Mus spretus (SPRET/Ei) using [32 P]-labeled CH7 as a probe revealed a restriction fragment length polymorphism for the enzyme TaqI. This polymorphism was used to follow the inheritance of the praja-1 gene using the (B6 X SPRET/Ei) X SPRET/Ei backcross panels (BSS) from The Jackson Laboratory Backcross DNA Panel Map Service (Rowe et al., Mammalian Genome 5:253-274, 1994). Linkage and order relative to other markers was determined by minimizing the number of multiple recombinants within each haplotype.

Northern blot analysis of Praja-1 expression: Northern blots containing 2 micrograms of poly(A)+ mRNA from mouse tissues (Clontech) were probed with [32 P]labeled CH7 antisense strand using Express Hyb hybridization solution (Clontech) at 68°C,

washed according to manufacturer's instructions, and subjected to autoradiography. A [32P]-labeled b-actin probe supplied with the Northern blots was used as a control to normalize RNA levels in each lane.

In vitro transcription/translation:

A transcription/translation-coupled rabbit reticulocyte lysate system (Promega) was used, as per manufacturer's instructions for [35S]methionine labeling. Clones of praja-1 in pGEM3Zf(-) plus a luciferase control clone were used with T7-RNA polymerase (sense direction). Each reaction comprised 12.5 ml rabbit reticulocyte lysate, 1 ml reaction buffer, 0.5 ml 1 mM amino acid mix minus methionine, 0.5 ml T7-RNA polymerase, and 20 units RNasin, all in 25 ml final volume. After a 90 min incubation at 300°C, products were lysed in SDS/mercaptoethanol treatment buffer and separated on a 10% SDS-polyacrylamide gel according to Laemmli, Nature 227:680-685 (1970). Proteins were electrophoretically transferred onto a BAS-NC membrane (Schleicher & Schuell) using a BioRad Trans-Blot apparatus according to manufacturer's instructions. Labeled products were visualized by autoradiography.

Results:

Isolation and sequence analysis of the novel gene, praja-1:

As part of the analysis of genes involved in liver development and function, we amplified the 3' end of a previously undescribed gene, CH7. We used the CH7 probe to screen a mouse embryonic cDNA library and isolated two overlapping clones, praja-1-5 and praja-1-6. Sequence analysis of the consensus overlap region revealed an open reading frame (ORF) of 424 amino acids, with a predicted size of 47.4 kD. Hydropathy analysis (Kyte et al. J. Mol. Biol. 157:105-132, 1982; not shown) shows that the translation product is highly hydrophilic, with no hydrophobic leader or membrane-spanning regions. The translation is also very acidic, with a pI of 4.6 and containing 17.7% acidic residues (Asp plus Glu). The

putative ATG start codon indicated in Figure 3 was selected because it is the upstream-most ATG that is in-frame with the ORF, and is preceded 21 bp upstream by a TAG stop codon. The context of this ATG, however, is only a weak fit to the consensus Kozak recognition sequence GCCACCatgG in that it does not have a purine at -3 nor a G at +4 (reviewed by Kozak, Genome 7:563-574, 1996). Sequence analysis of the amino acid translation revealed the presence of a COOH-terminal RING-H2 motif, which is a zinc finger variant (Freemont, supra). Figure 13 shows an alignment of the RING-H2 motif of praja-1 with those of several other RING-H2 containing proteins.

Linkage analysis places Praja-1 on mouse chromosome X:

A restriction fragment length polymorphism for praja-1 was identified using CH7 as a probe on a Southern blot containing DNA from the two parental strains digested with several restriction enzymes (TaqI, BglII, EcoRI, EcoRV, HindIII, HincII, KpnI, PstI). For every enzyme used, C57Bl6/J had only a single restriction fragment, while two fragments were always observed within the SPRET/Ei lane. A polymorphism obtained using TaqI was used to type the inheritance of the C57Bl6/J allele in the BSS panel. There are two Spretus bands S1 and S2 and one C57Bl6/J band B1. After comparison of the praja-1 genotypes to other genes typed within the database, it was determined that praja-1 maps to mouse chromosome X at about the 36 cM position (Fig. 14). The S1 band is the praja-1 allele on X chromosome of SPRET/Ei. The S2 TaqI fragment appears in every backcross animal. Since all males from the backcross contain this allele, it is not localized to the X chromosome. Since females also have the S2 band, it is not Y-linked. Therefore S2 is an autosomal locus that contains sequence homology to the praja-1 probe sequence. Other genes mapping to this general region include moesin (Msn), androgen receptor (Ar), interleukin-2 receptor gamma (Il2rg), X-linked zinc finger protein (Zfx), and tabby (Ta). This area is also 1.1 +/- 1.1 cM from the Xist locus. Further studies are needed to

determine if praja-1 is not expressed on inactivated X-chromosomes and if it plays a role in X-inactivation. The syntony and conserved gene order between mouse and human X chromosomes (Herman et al., Genome 6:S317-S330, 1996) allows comparison with human disease genes in the region. Human diseases in this region with mesodermal involvement include anhidrotic ectoderm dysplasia (eda) and sideroblastic anemia with spinocerebellar ataxia (asat).

In vitro expression produces a protein product larger than the predicted size. An autoradiogram of the in vitro transcription/translation products of clones praja-1-5 and praja-1-6 showed that only praja-1-5 produced a significant product. The product, which ran as two closely spaced bands of $M_r = 55.6$ and 56.9 kD, is larger than the predicted ORF size of 47.4 kD. One possible explanation is that the expression product is very acidic, and acidic proteins such as granins are known to give anomalously high M_r on SDS-PAGE (Huttner et al., Trends Biol. Sci. 16:27-30, 1991). The presence of two products suggests translation initiation at a second, internal ATG codon, such as at Met- 19 (Fig. 3).

Praja-1 transcripts are present in embryonic and in mouse tissues. Northern blot analysis of RNA from adult mouse showed expression of 3.1, 2.6, and 2.1 kb transcripts in liver, brain, and kidney, and an additional 2.3 kb transcript in testis. The praja-1 protein is unlikely to be a membrane receptor, since it lacks a hydrophobic transmembrane domain. The uniform hydrophilicity suggests a soluble protein. The praja-1 RING-H2 motif is shown aligned with those from several other proteins in Figure 13. RING fingers are generally thought to function in protein-protein interactions (Borden et al., Curr. Opinion Struct. Biol. 6:395-401, 1996; Saurin et al., Trends Biochem. Sci. 96:208-214, 1996). To cite a specific example, if either of the two cysteines that comprise the Zn^{++} binding site of the RING finger of acute promyelocytic leukemia protooncprotein PML are mutagenized, then the nuclear multiprotein complex, or so-called nuclear bodies, fail to occur (Borden et al., EMBO J.

14:1532-1541, 1995). The authors conclude that the PNM RING domain, and probably other RING finger domains, are involved in protein-protein interactions.

In praja-1, aside from the RING-H2 finger, the stretch of thirty-four COOH-terminal amino acids just past this motif (Fig. 3) is especially rich in proline residues (17.6%); and, as stated, the protein in general is very acidic. Proline-rich domains are found in several mammalian transcription factors, such as that at the COOH-terminus of transcription factor CTF, and proline-rich regions and also acidic regions are likely to function in contacting other proteins (Mitchell et al., Science 245:371-378, 1989). A BLAST search of the proline-rich COOH-terminus revealed no significant matches to any protein in the available databases, however, when considering the praja-1 sequence as a whole, the rat Neurodap1 gene has the highest similarity; the alignment is presented in Figure 15.

Neurodap1 is expressed abundantly in rat brain, with much smaller amounts in heart and skeletal muscle. Though praja-1 likewise shows greatest expression in brain, unlike Neurodap1 it also expresses in liver and kidney. The subcellular localization of Neurodap1 was shown to be concentrated around the endoplasmic reticulum (ER) and golgi of the cerebral cortex and facial nucleus, and especially in the postsynaptic density region of axosomatic synapses (Nakayama et al., supra). Based on its subcellular localization, plus the presence of the RING-H2 finger, the authors concluded that Neurodap1 is probably linked to the secretory or protein sorting. Praja-1 does differ from Neurodap1 in several respects, however. In addition to being expressed in some different tissues than Neurodap1, praja-1 encodes for a product that is smaller (47.4 kD, based on the composite of the clones in Fig. 3) vs. 77.9 kD for Neurodap1. The difference in size is at the N-terminus of the proteins (Fig. 15). The largest transcript we observed for praja-1 was 3.1 kb, whereas Neurodap1 exists as a single 4.8 kb transcript on Northern blots of rat brain mRNA.

In light of the fact that BRCA1, which possesses a RING

finger, has an acidic pI, and is a secretory protein, also has properties of the granin family of proteins (Jensen et al., Nature Genet. 12:303-308, 1996), we examined praja-1 for a granin signature. We found no region in the praja-1 translation that gave a perfect match to the consensus E[N/S]LX[A/D]X[D/E]XEL, though two regions matched five of the seven conserved residues. We were also unable to demonstrate the presence of clear coiled-coils, which are present in BRCA1 and proteins with the previously-mentioned tripartite structures. In these respects, praja-1 is more similar to Neurodap1 than to proteins such as BRCA1. Also, though the RING-H2 finger in praja-1 shows much similarity to that from the *D. melanogaster goliath* (gl) protein (Fig. 13), the goliath protein possesses an alkaline pI (8.9) and no sequence similarity to praja-1 outside of the RING-H2 finger. The RING-H2 motif plus acidic and proline-rich regions, and similarity to Neurodap1; leads to the conclusion that praja-1 is involved in protein-protein interactions, possibly in a protein sorting or secretory pathway.

EXAMPLE 2

In accordance with the present invention, investigations were made with regard to the induction of differentiation in liver tissues in order to isolate and identify early developing liver proteins for use in therapies involving the liver and liver functions. In the developing fetus, inductive interactions, intercellular communication and the establishment of cell polarity are critical for growth and patterning during development. However, the precise mechanisms by which these effect hepatocyte differentiation or liver development have not previously been elucidated. Mammalian liver development was first recognized to be established through a specific sequence of interactions between mesenchymal and endodermal embryonic tissues. At 9.5 days of mouse gestation, upon signaling from the cardiac mesenchyme, endodermal cells from the liver diverticulum proliferate and migrate into the surrounding

septum transversum. This specific area of loose mesenchyme in turn differentiates into hepatic mesenchyme and a liver bud is finally recognizable microscopically at about 10.5 days of gestation. This hepatic mesenchyme is continually responsible for the hepatocyte proliferation which then proceeds throughout embryonic life (Le Douarin, Med. Biol. 53:427-455, 1975). Albumin transcription can be detected as early as at day 9.5 (Cascio et al., Development 113:217-225, 1991), implying that hepatocyte differentiation begins when hepatic endoderm comes into contact with cardiac mesoderm. As a first step towards the analyses of signal transduction pathways regulating such a restricted pattern of gene expression, molecular markers as well as regulatory genes are required to identify the interactions required for liver development.

The dissection of gene regulatory pathways in the liver has led to the identification and characterization of transcriptional activators, C/EBP, DBP, LFB 1/HNF 1, 3 and 4 (Johnson, Cell Growth Differ. 1:47-52, 1990; Kuo et al., Development 109:473-481, 1990; Frain et al., Cell 59:145-157, 1989), of liver specific genes, such as α -fetoprotein and albumin (Tilghman, Oxford Surveys on Eukaryotic Genes, Oxford University Press, 1985). Yet, with the exception of HNF4, 3 α and β (Ang et al. Development 119:1301-1315 (1991) and Cell 78:561-574, 1994), none of the above have been found to play a definitive role in determining cell-lineage and regional specification of the developing liver. The small volume of liver buds (approximately 4×10^{-2} mm³) yields even smaller quantities of proteins, DNA and messenger RNA thus making the molecular analysis of liver development difficult. Therefore, the construction of early embryonic liver cDNA libraries, and performing subtractive hybridization still remains the most plausible and comprehensive method of obtaining an unbiased catalogue of genes required during early mouse liver development (see Harrison et al., Development 121:2479-2489, 1995).

The isolation of markers would provide further insight into identifying transcriptional activators and growth factors involved in such a restricted pattern of gene expression, and eventually provide an approach to identifying signal transduction pathways involved in hepatocyte differentiation. In some cases, these pathways have been characters as in patterning and axis formation of the vertebrate head and body (Oliver et al., Development 121:693-705, 1995; Kessel et al., Science 249:374-379, 1990). For example, in Xenopus, a network involving brachyury, activin and wnt-related genes, is responsible for mesoderm induction, somitogenesis, myogenic and sclerotomal differentiation (see, e.g., Wilkinson et al., Nature 343:657-659 (1990); Herrmann et al., Development 113:913-917 (1991); Green et al., Trends Genet. 7:245-250 (1991); Sokol et al., Cell 67:741-752 (1991); Smith et al., Cell 67:753-767 (1991), and dorsal ventral axis formation results from Xgsk-3 (the Xenopus homologue of Drosophila zw3/shaggy) phosphorylating its Xenopus homologue of armadillo, β catenin thus regulating the level of β catenin available for dorsal axis formation. However, there are no available molecular markers nor pathways which chamctedw either earlier liver development, nor its crucial mesodermal component.

In accordance with the present invention, it has been possible to identify and characterize such molecular markers and possible inductive transcripts for liver development. As set forth below, the characterization of the elf protein is described, and the expression of this protein may mark the separate components of liver development. The "bottom up" approach with regard to this charatcerization in general has led to the identification of a totally unexpected group of genes, and in particular, this is described with regard to the elf protein, which is probably involved in playing a role in establishing cell polarity by interactions at the surface membrane.

Characterization of cDNA libraries:

The four stages in liver development (e10, e11, e12, and e14, where e=embryonic) are defined developmental time points from undifferentiated mesodermal/endodermal cells to a well developed and differentiated fetal liver. A change in cell polarity occurs at e9-10. At e10.5-11, invasion and migration of endodermal cells into surrounding mesenchyme occurs; at e11.5-12, pseudolobule formation, cords of hepatocytes form together with early sinusoids. cDNA libraries representing these stages would therefore, represent "captured" mRNA species expressed in greater abundance during crucial time periods for hepatocyte formation, enabling their isolation and providing a method for analyzing the changing pattern of gene expression during liver development.

Libraries containing 5.0×10^6 - 4.1×10^7 independent clones were generated from the largest cDNA fractions. Current estimates demonstrate that a library containing 5.0×10^5 clones (Sambrook et al, Molecular cloning, a laboratory manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1989) is a representative library with a 99% probability that rare transcripts (less than ten copies per cell) are present. Our libraries are therefore likely to be truly representative of their respective mRNA species for that stage.

Qualitative and developmental profiles of the libraries:

These were obtained, utilizing genes, such as IGF-II, IGFBP-2, IGF1, C/EBP, HNF/LFBI known to be expressed at different time points in developing liver. The data in Table 2 below demonstrate that IGF-I was not detected in any of the embryonic libraries, while IGF-H was detected in the e10.0 and e12.5 libraries (3 at e10.0 and 4 at e 12.5). IGF-II was not detected in the adult liver library. Interestingly, BP-2 clone frequencies are similar to IGF-II in the early e6.5, e7.5 and e8.5 libraries (data not shown), but in the liver cDNA libraries the clone frequencies differed, for BP-2 only one clone per 100,000 being detected at e10.0 and e11.5, while 7 were detected in the adult liver cDNA library compared to the

greater numbers for IGF-II. This implied that its temporal and spatial expression in the embryo and fetus is different from IGF-II and this was subsequently confirmed by in situ studies.

HNF1/LFB I detected in the e12.5 library was suddenly detected at day 11.5 and 12.5 in low abundance (2 clones/100,000 at e11.5 and 5 at e12.5), confirming that while it is expressed, its level also may be regulated, albeit downward, in embryonic stages. Lastly, mouse β -Actin was used as a reference: all seven libraries had similar β -Actin frequencies from 120-300/100,000 clones which is considered representative of such embryonic libraries.

Identification of stage specific clones by subtractive methods:

Two subtracted libraries were then constructed as previously described, comprising 64 clones (e11.5-12.5), and 174 clones (e10.5-11.5). Further characterization of these clones was carried out by Southern hybridization, sequencing, Northern blot analysis, Zoo blot analysis, and in vitro fertilization of protein. Using Southern blotting, thirty-four clones were shown to be stage specific and not containing mitochondrial, ribosomal and globin sequences, and further analysis was carried out on elf.

Identification and developmental regulation of elf transcripts:

Elf mRNA in tissues from mid- gestational embryos were analyzed, and tissues were dissected from day 11 onwards since it was at this stage that discrete hepatic, cardiac and other tissues could be dissected with ease, and the subsequent RNA isolated was of good quality. Using a 32P-labeled 1.1 Kb insert representing elf, the specificity of the developmental changes in the steady state levels of elf was evaluated by also measuring the relative levels of β -Actin. This revealed a 2.1 Kb transcript at high stringency washes. Scanning densitometry of the respective bands revealed that maximal expression of elf occurred in liver and heart, less so in other tissues but

specifically on day 11, and in 12.5, 14.5 in decreasing abundance (when Northern blots were developed 1-2 months later).

Sequence analysis of elf:

After subtraction hybridization, one stage specific clone was analyzed in detail: sc32. The initial libraries were then screened at high stringency (0.2x SSC, 60°), to obtain overlapping clones for sc32. Positives were picked, and after in vivo excision (Stratagene) into Bluescript, these were sequenced using the dideoxy chain termination method using oligonucleotides corresponding to previously determined sequence. Of the seven clones picked, three were found to be overlapping to sc32 and included sequence encoding elf (Figures 16a and 16b). Confirmation of the identity of the clones and elf was carried out by Northern blot analysis of mouse embryonic tissues. In the case of elf, this gave rise to the same initial 2.1 Kb transcript with sc32 as a probe. A start codon was not present suggesting that we had not cloned the 5' end of the cDNA. However, the northern blot showed a 2.1 Kb transcript, thus suggesting that we had cloned complete elf and this probably represented a spliced form of β -fodrin. The authenticity of the 3' end of the elf sequence was confirmed by the comparison of the elf sequence with the expressed sequence tags (EST) database. Although no mouse ESTs for elf sequence were found, three different human EST clones were found to span the region of unique last 100 nt and the 5' adjacent sequence, suggesting the existence of elf homolog in human cells (see Figures 16a and 16b).

Prior sequence analysis has shown elf to bear 80% identity to β -fodrin, a non erythroid β -spectrin. Our sequence to elf is located between domains II and III of the β -spectrins. Domain II comprises 17 repeats of a 106 amino-acid motif and an ankyrin binding domain (Figure 16a). The ankyrin binding domain is required for the correct subcellular localization of adducin, ankyrin and the Na⁺,K⁺ ATPase, without which cell

morphology is disrupted. Domain II comprises a C terminal domain which contains varying numbers of residues (52-265) in alternatively spliced forms giving rise to tissue specific expression (Hu et al., J. Biol. Chem. 267:18715-18722, 1992), as well as the PH domain.

In situ localization of elf:

In situ hybridization confirmed elf expression in 11.5 heart and liver and determined its expression pattern during earlier liver development, using elf sense probes and alpha fetoprotein antisense probes as controls. The hepatic diverticulum, which originates at the foregut-midgut junction, begins to grow into the septum transversum at the 9th day of gestation (13-20 somite stage). Between days 10.5 to 11.0 p.c., a considerable degree of differentiation is seen in this primitive liver. The liver enlarges substantially over this period: the increase in the overall volume being due to the invasion of the mesenchyme of the septum transversum by the hepatic cords, and the initiation of hematopoietic activity in this organ. At day 9.5, a strong labeling of elf becomes apparent in the cardiac silhouette: the pattern appears to be trabecular, including the wall of the cardiac anlage. A section of the cephalad chamber (sino-atrial chamber) wall also bears a high intensity of elf expression. The surrounding tissue, particularly the caudal liver bud region does not show the presence of silver grains. At the next stage, day 10.5, silver grains clearly highlight the developing liver, which appears as a horizontal oriented structure (L) in this section.

At this stage, the signaling is weakening in the developing heart tissue. The surrounding tissues are remarkable for the absence of silver grains. At day 11.5, a strong labeling becomes apparent in the liver, which is larger in size. The heart shows an extremely weak signal: silver grains being visible in only a single streak posteriorly. At this stage, elf expression also appears in the umbilical cord. As a control, in addition to sense probes, a riboprobe to

α -fetoprotein outlines the developing embryonic liver at day 11-12.

A comparison of the day 9.5 and 10.5 embryos demonstrates a clear temporal and spatial gradient of maximal tissue staining with silver grains representing elf riboprobe: the temporal gradient of a rise and fall of elf expression in the heart may be inferred from the strong staining in the developing heart at day 9.5 followed by a weaker staining at the next stage (day 10.5). Simultaneously, liver expression increases. The spatial gradient is apparent from the developed patterns of these tissues which showed that silver grains increase in density as one moves from the developing heart to the liver: at day 10.5, antisense RNA probes from elf cDNA hybridized specifically to 9.5 day cardiac mesenchymal tissue; expression at day 10.5 being restricted to cardiac and hepatic tissue; elf expression finally being restricted to the liver in later 11.5 day embryos. Of note, elf expression was seen in embryonic livers at later stages (days 12.5, 14.5 p.c.), but only in decreasing abundance: message being detected in these later stages when Northern and in-situ were developed a considerable time later. Elf sense probes did not hybridize to any tissues.

Alpha fetoprotein antisense RNA probes hybridized specifically to 11.5, 12.5, 14.5 embryonic mouse liver tissue, in agreement with previous studies of mRNA isolated from embryonic liver samples (Tilghman et al., P.N.A.S. 79-5254-5257, 1982). The earliest stage of detection of α -fetoprotein mRNA by in situ hybridization was at 10.5-11.0 days of gestation. Similar experiments with albumin mRNA (Cascio et al., Development 113:217-225, 1991) have shown it to be expressed at 9.5d in clusters of cells arising from foregut epithelium and in cords of cells beginning to invade the septum transversum. In the experiments with α -fetoprotein, the liver was labeled at all subsequent stages (day 11 onwards), and, upon histological examination appeared to occur primarily in

the endothelial cells. Hematopoietic cells appeared refractile but did not contain the hybridization grains that were visible over the α -fetoprotein positive cells.

ELF mRNA distribution in mesodermal tissues versus Alpha fetoprotein mRNA in endodermal tissue:

Since Northern analysis revealed elf expression to occur in day 11.5 heart and liver tissue, we investigated whether elf expression was restricted specifically to mesodermal tissue from the heart and the liver and compared this to the endothelial expression of α -fetoprotein. Three main regions of mesoderm can be discriminated in the developing embryo: dorsal (somitic), intermediate, and lateral. Lateral plate mesoderm comprises somatic (pleura, pericardium, peritoneum and limb bud), and splanchnic (heart- epicardium, myocardium, connective tissue and smooth muscles of viscera and blood vessels, hemangioblastic tissue, adrenal cortex and spleen). Regarding the developing heart, at day 9 (13-20 somites), this is seen to beat regularly and strongly. At this stage, the heart appears to be the only region within the embryo where the endothelial elements of the circulation are surrounded by a vessel wall. The walls of the common ventricular chamber as well as the common atrial chamber show an increasing degree of trabeculation. Of note, the space between the endothelial and myocardial elements is filled with loose mesenchyme called cardiac jelly. In situ hybridization of days 9 and 10 embryonic heart tissue using elf antisense riboprobes demonstrated high levels of labeling to both the atrial and ventricular regions.

Hepatic mesenchyme also originates from lateral plate mesoderm. The septum transversum part of the hepatic mesenchyme originates from the splanchnic mesoderm of the precardiac area and this is thought to be responsible for the subsequent differentiation of hepatocytes. However, tissue explant experiments have shown that all derivatives of the lateral plate can replace hepatic mesenchyme for these later

events. While these initial experiments have demonstrated migrating endoderm must interact with mesenchyme for the former to differentiate into hepatocytes (Le Douarin, 1975; Houssaint, Cell Differ. 9:269-279, 1980), more recent studies investigating albumin mRNA expression as an indicator of hepatocyte differentiation, have confirmed these features: initial expression of albumin mRNA occurs during the invasion of the septum transversum, when the hepatic precursor cells clearly contact cardiac mesenchymal tissue. Similarly, primer extension analysis of albumin transcription has revealed the start site of transcription to occur at day 10.5 with a 15-20 fold increase in albumin mRNA upon liver organ formation by day 12.5. In our experiments using α -fetoprotein as a marker for differentiated hepatocytes, it was clear under high magnification, that while α -fetoprotein expression is restricted to the later endodermal component of liver development, elf expression seems to occur in the loosely organized, lighter staining mesenchymal cells - initially cardiac mesenchyme (at day 9.5), then in both cardiac and hepatic tissue (at day 10.5) and then restricted to liver tissue (day 11.5 onwards; elf expression then decreasing upon liver formation. Examination of the later histological sections (days 11 onwards) showed a diffuse distribution of grains. The resolution that was attained did not allow one to draw a firm conclusion about the identity of the hybridizing cells, although it seemed that the hybridization signal with elf was localized in the perisinusoidal cells, but not in the hepatocytes.

Distribution of elf RNA in Adult tissues, conservation in evolution:

Further characterization of elf has involved RNA analysis of adult mouse tissues. Elf hybridizes to adult liver, kidney and testis as a 2.1 Kb transcript in liver and kidney and a 2.6 Kb transcript in adult testis, in very low abundance. Genomic analysis of elf DNA from human, monkey, rat, mouse, dog, cow,

rabbit, chicken and yeast indicates that elf is conserved across the species, being represented in all except rabbit DNA (Fig. 18).

In vitro transcription and translation of elf, the latter using nuclease-treated rabbit reticulocyte lysate (promega), has revealed a 34 Kd protein, which is as predicted by the elf insert size and indicating that this insert is in frame for the coding sequence for a specific protein (Figure 17).

Embryonic liver explants cultures:

One of the goals of the investigations in conjunction with the present invention was to establish a functional assay for determining the developmental roles of elf and ss3 in liver formation. Mouse embryonic liver explants were cultured in our laboratory, in order to overcome the dissection and analysis of extremely small tissue sections at day 10-10.5 when the liver bud is 0.2 mm. When cultured in the complete absence of mesodermal derivatives, hepatic endoderm deteriorates rapidly.

Only 2 out of 15 such liver explants survived. Hematoxylin and eosin staining showed a necrotic endoderm with no apparent signs of hepatic differentiation. When associated with the surrounding mesoderm particularly cardiac mesoderm (en bloc dissections), the endodermal cells had proliferated and invaded the mesoderm strands. Hepatocytes were seen to be organized in cords separated by sinusoids with pseudo-lobule formation. All 15 out of 15 cultures from en bloc dissections were completely viable. These studies confirm prior explant studies demonstrating the necessity of surrounding mesoderm for liver formation. Semi-quantitative RT-PCR analyses of elf, other clones ss3, 145, HNF 3 β with GAPDH and α -fetoprotein as controls demonstrate increased expression during mesodermal - endodermal interactions.

Early experiments in chick embryos (Douarin, 1975, supra) have demonstrated that at the primitive streak stage, the prospective hepatic area is localized in the middle and in the lateral areas anterior to Hensen's node. At the head process

stage, prospective liver areas coincide with cardiac areas, being concentrated in bilateral areas extending from the tip of the head process to an area slightly behind the primitive pit.

Potential liver areas were tested by transplantation of pieces of tissue on the chorioallantoic membrane; liver differentiation in such explants was dependent upon the presence of cardiac tissue: no liver tissue was found without cardiac cells in the vicinity, whereas some grafts contained heart tissue without liver. After gastrulation is completed, it is during the somitic stage that the liver and heart segregate partially - the presumptive cardiac mesenchyme migrates anteriorly and venally into the cardiac fold, the prospective myocardial cells becoming incorporated in the heart anlage. Another series of experiments using carbon particle labeling, radiodestruction and coelomic transplantation of pieces of blastoderm showed liver endodermal and mesodermal areas which are superimposed during the early embryonic stages evolve differently later on.

Tissue explant studies have revealed that in normal liver development, hepatocyte differentiation and the formation of liver lobes is entirely dependent upon the mesodermal component which then becomes progressively colonized by the growing endoderm hepatic cords (see Douarin, 1975, supra). These stimulating properties of the cardiac, and then, hepatic mesenchyme have been demonstrated to begin at the 5 somite stage and last throughout embryonic life. The findings set forth herein show that *elf* is expressed in early cardiac mesoderm, with subsequent expression being limited to hepatic mesoderm, revealing this to be a novel marker for the mesodermal component of liver development. Of note, in normal development, pure liver mesenchyme is never observed. That these explant studies have demonstrated expression of *elf*, indicates that *elf* protein will be useful in identifying and studying interactions between mesoderm and foregut endoderm.

Summary of events during hepatocyte formation indicating a role for *elf*:

Embryonic	endodermal cell hypertrophy	
Stage		<i>elf</i>
day 9.5	change in cell polarity	expression
day 10.5	invasion and migration into surrounding mesenchyme	
day 11.5	pseudolobule formation, cords of hepatocytes, early sinusoidal formation	
day 14.5	hematopoietic foci and fully differentiated fetal hepatocytes	

Sequence analysis has shown *elf* to bear 80% identity to β -fodrin, a non erythroid β -spectrin. β -spectrins have been implicated in numerous functions including the maintenance of cell surface polarity of cells (Nelson et al., J. Cell. Biol. 108:893-902, 1989); the maintenance of cell-cell junctions (Thomas et al., Development 120:2039-2050, 1994, Luna et al., Science 258:955-964, 1992); β -spectrins contain binding sites for other proteins, such as ankyrin and actin (Hu et al., J. Biol. Chem. 267:18715-18722, 1992; Speicher et al., Nature 311:177-180, 1984). Smaller isoforms β -spectrins have been well described. For instance, a 4.0 Kb muscle tissue transcript is thought to encode a previously reported β -spectrin from clustered acetylcholine receptors. Similarly for *elf*, the missing domains may be replaced through alternate exon usage to generate proteins with unique functions. A function for *elf* thus appears to be in the assembly and maintenance of specific domains on the cell surface - towards establishing hepatocyte polarity and thus differentiation.

Spectrins have also been shown to be conserved throughout evolution and are developmentally regulated. These results demonstrate that in keeping with brain β -spectrin (β -G spectrin), *elf* is also expressed in a tissue and stage specific manner and is conserved throughout evolution (Hu et al., J.

Biol. Chem. 267:18715-18722, 1992; Zimmer et al., Brain Res. 594-75-88, 1992; Leto et al, Mol. Cell Biol. 8:1-9, 1988). Elf expression occurs in a gradient-like manner and close examination of Brain β -G spectrin has demonstrated similar gradient patterns, suggesting that a sudden on-off phenomenon at specific time points is simplistic. That elf is maximally expressed at day 10-11 suggests that it has an important function at this time, which continues, although to a lesser extent, with the later stages. For instance, it is conceivable that elf by conferring cell polarity mark the first overt sign of hepatocyte differentiation. Therefore, like *Drosophila* β -H spectrin, elf may play a role in facilitating a "velcro-like" joining of neighboring cell membranes as they extend (Thomas et al., Development 120:2039-2050, 1994). In this way elf may mark the polarization of the surrounding mesodermal cells, enabling foregut endodermal cells to invade this area and differentiate into hepatocytes. Molecular markers have been invaluable in the dissection of inductive events in embryological studies (New et al., Curr. Opin. Genet. Dev. 1:196-203, 1991; Sive, Genes Dev. 7:1-12, 1993). For instance, in *Xenopus*, Epi 1, an antibody specific for epidermis, has been used to elucidate the role of the blastopore lip in the neural induction process (Savage et al., Dev. Biol. 133:157-168, 1989). Similarly activins (regulating keratin) (Asashima et al., P.N.A.S. 88:6511-6514, 1991), vg-1 (Thomsen et al., Cell 63:485-493, 1990) and other genes belonging to the TGF- β family, as well as wnt and bFGF families represent components of the cascade leading to the commitment to particular mesodermal fate. For instance, vg-1, originally isolated by differential screening is to cells inducing embryonic mesoderm, the posttranslational processing of Vg-I precursor protein on the future side of the embryo being a key step in generating dorsal mesoderm and body axis in *Xenopus* (Thomsen et al., Cell 74:433-441, 1993). Similarly, in isolating putative inductive agents required for liver formation, a key step is the

identification of mRNAs localized to cardiac/liver mesenchyme: elf and its regulatory genes will help to elucidate this area.

More recently, cell-cell interactions have been shown to be important for several cell fate decisions. In C. elegans for instance, *lin-12* and *glp-1* have been shown to encode transmembrane proteins mediating intracellular communication, and are required for the specification of several anterior fates. In *Drosophila*, the establishment of secondary epithelia which are the result of a mesenchymal-epithelial transition, is thought to be dependent upon two separate adhesions systems: direct interactions between the developing midgut endoderm and the visceral mesoderm on one hand and, adhesive interactions between the epithelial cells themselves on the other. While the latter cell-cell interaction is thought to be controlled by shotgun, control of apicobasal polarity is thought to be caused by genes such as *cnaab*s and *stardust* (Tepass et al, Cell 61:787-799, 1990). Although it is known that the biogenesis of cell surface polarity in hepatocyte formation is an early event, implying that the mechanisms for sorting plasma membrane molecules are functional at an early point, genes involved in cell signaling leading to cell fate in liver development have not been defined to date. The identification of such genes would give tremendous insight into the cell-cell interactions involved in foregut endodermal cell migration and subsequent morphogenesis of the liver as an organ. These studies establish the principle that specific mesoderm mRNAs are localized in a way that guarantees their subsequent segregation to specific mesodermal tissue, in this case the presumed mesodermal component of the liver as demonstrated by embryonic explant studies (Le Douarin, 1975).

Cloning and sequencing of elf:

All embryonic liver was obtained from matings of random-bred ICR mice (Harlan). The plug date was designated as Day 0 and embryos collected at days 10.0, 11.5 and 12.5 p.c.; these were staged by morphological criteria (Theiler, The House

Mouse, New York: Springer-Verlag, 1989). The livers were dissected, pooled and lysed. To prepare cDNA libraries, RNA was isolated (Chomczynski et al., Analyt. Biochem. 162:156-159, 1987) and poly(A)+ RNA selected using oligo(dT)-cellulose (Collaborative Research Type 3). 1 to 5 mg of poly(A)+RNA were used in the preparation of oligo(dT)- primed cDNA libraries. cDNA library construction of days 11.5 and 12.5 embryonic liver was carried out by conventional techniques (Gubler et al., Gene 25:263-269, 1983), and the day 10.0 and adult mouse liver using the Stratagene Unizap cDNA library kit. Two subtracted libraries were then constructed (Schweinfest et al., Genet. Anal. Tech. Appl. 7:64-70, 1990). The resulting subtracted libraries comprised 64 clones (11.5-12.5), and 110 clones (10.5-11.5). The process involved: (a) Biotinylation: fifty micrograms of cDNA from day 12.5 liver library at 10mg/ml were biotinylated in HE buffer (10mM Hepes, pH 7.5, 1mM EDTA, Clontech Labs.); (b) Subtraction was then done by the streptavidin-phenol extraction: the streptavidin-biotin hybrid duplexes represent common gene products which selectively partition into the phenol interface, leaving the unique, subtracted single stranded cDNA in the aqueous phase. After synthesis of second strand DNA and overnight precipitation, one tenth of the DNA was used to transform competent XL Blue cells. Transformation using all the subtracted DNA led to the identification of 174 recombinant colonies. Purification of bacteriophages, preparation of DNA were carried out by the stratagene in vivo excision protocol. Plasmid DNA was sequenced using 77 DNA polymerase (Sanger et al., J. Mol. Biol. 143:161-178, 1980).

Sequence analysis:

The NCBI non-redundant (nr) and EST databases were searched using the blastp2 and blastn2 programs, which permit gapped alignments (Altschul et al., Methods in Enzymology 256:460-480, 1996), with the default parameters and elf protein or nucleotide sequences as queries.

RNA preparation and analysis:

Embryos were collected at day 10.0, 11.5 and 12.5 p.c. Embryonic livers were dissected in Dulbecco's modified Eagle's medium (high glucose) and 20 mM Hepes pH 7.3. The livers for the specific stages were pooled and total RNA isolated (Chomczynski et al., supra). 10 micrograms of RNA were electrophoresed on a 1% formaldehyde gel and transferred onto Hi-bond nylon membrane (Amersham) using standard procedures (Sambrook et al., 1989, supra). Radioactive, ^{32}P -labeled probes were synthesized by random primer methods (Feinberg et al., Analyt. Biochem. 137:266-267, 1984) and hybridized to the Nylon filters. Filters were washed at high stringency with a final wash in 0.2x SSC (30 mM NaCl, 3mM sodium citrate, pH 7.4) 0.5% Sodium Dodecyl Sulfate at 65°C for 60 minutes. Filters for each probe were stripped and rehybridized with other probes to confirm that no cross hybridization signals were obtained under initial screening conditions. These filters were then autoradiographed with intensifying screens at -70°C.

In Situ Analysis:

In situ analysis was performed for elf (Cox et al., Dev. Bio. 100:197-206, 1989). The RNA probes were synthesized and labeled with ^{35}S -UTP (400 Ci/mmol) via the T7 or SP6 promoter for RNA polymerase. Sense or antisense probes were added to the appropriate sections, mounted, sealed with rubber cement and incubated at 50°C overnight. After incubation, sections were washed with 50% formamide/5xSSC/10mM DTT (50°C; 2X 30 min.) followed by 4x SSC/TE, incubated with RNase A (20 mg/ml) and RNase TI (500 U/ml; 37°C 30 min), rinsed again with 4x SSC/TE (37°C, 30 min), twice 2x SSC (25°C, 15 min), twice in 0.1x SSC (25°C, 15 min), dehydrated with an ethanol series (containing 0.3 M ammonium acetate) and air dried. For autoradiography, slides were dipped in NTB 2 emulsion diluted 1:1 with 2% glycerol in water and dried. Exposure times were from @ weeks

to four months. The emulsion was developed according to manufacturer's directions.

In Vitro Translation of elf:

Bluescript containing elf was transcribed with T7 RNA polymerase using the in Vitro Eukaryotic Translation kit and MCAP mRNA Capping kit (Stratagene). The RNA transcript was translated in vitro into protein for 90 minutes in the presence of [³⁵S]methionine using nuclease-treated rabbit reticulocyte lysate (Promega) and run on 4% denaturing polyacrylamide gels.

Liver explant cultures:

Mouse embryos were obtained from Harlan ICR mice. The age of the embryos was determined by days post appearance of the vaginal plug (day 0). The embryos were further characterized by the number of somites. Isolation of mouse hepatic endoderm, liver buds and mesoderm (en bloc dissection) was as follows: during the 10th day of gestation, the liver bud becomes evident as a thickening of the ventral wall of the foregut, near the origin of the yolk stalk. This ventral endoderm was then either taken alone and cultured, or alternatively with the surrounding mesoderm: the portion of the embryo between the otocyst and the umbilical region. Organ culture: Embryos were placed into nucleopore filters in a humid chamber as described (Houssaint, 1980, supra) and cultured for 48 hours or 96 hours.

Microscopy: The explants were fixed as in the in situ hybridization protocols, and RNA isolated as described above. 7mm sections were stained with hematoxylin, eosin and periodic acid schiff (PAS) for glycogen, an indicator of differentiated hepatocytes. For RNA analysis, semiquantitative RT-PCR was performed as described in Fig. 19 and 20. The invention contemplates the use of such liver explant culture for tissue engineered composites as a form of liver restoration therapy.

Immunohistochemical characterization:

Antibody to a peptide corresponding to amino acids 145-157

(CLRRKYRSREQPQS) of prajal (COVANCE), was used for immunohistochemical localization in liver explant cultures. Embryos were fixed and embedded into paraffin, sectioned and immunostained using indirect immunohistochemistry according to protocols routinely used (Schevach, Current protocols in immunology, Green Publishing Associates and Wiley Interscience, 1991). 8 μ m sections were deparaffinized in xylene, the tissue rehydrated in graded alcohols, and rinsed in PBS. The sections were initially treated with a protease (0.1% Trypsin in 0.05 M PBS) and incubated at 37°C for 30 min. Endogenous peroxide was then removed using 3% hydrogen peroxide. Sections were blocked in PBS containing 5% goat serum for 30 min. at room temperature. Sections were then incubated overnight at 4°C in a Humidor with diluted rabbit anti-mouse antibody directed against the PRAJA1 peptide. all further steps were done at room temperature. Six 5 min. rinses with PBS-S were performed after each successive step. After incubation in the primary antisera, slides were washed six times for 5 min. in 1 x PBS at room temperature. Sections were incubated with a second antibody (diluted in 0.05 M PBS in 1% serum) for 30 min. at room temperature.

After rinses the substrate was added as follows: Insoluble Peroxidase substrate DAB (Sigma Fast). 100-150 microliter substrate solution was added to cover the entire tissue on the slide. Color development was monitored under microscope. After rinsing in distilled water for 5 min, staining was performed with Harris hematoxylin solution modified (Sigma) for 1 min., followed by a rinse in distilled water for 5 min. Sections were dehydrated by passage through distilled water, then graded alcohol concentrations and finally xylene. Coverslips were mounted using DPX (Fluka labs) or Permunt (Fischer scientific), before observation. For the negative controls only the primary antibody diluting solution was added, without any antibody.

Generation of Antibodies:

Peptide-specific rabbit anti-mouse polyclonal antibodies to sequences in the N-terminal and C-terminal of ELF3 were generated as described in Porter et al., J. Cell. Biol. 117:997-1005 (1992). The sequences of the synthetic peptides were ELQRT SSVSG PLS (residues 2 to 14 at N-terminus) for the preparation of EL-1, and FNSRR TASDH SWSGM (residues 2140 to 2154 at C-terminus) for the preparation of EL-2. IgG was isolated from antisera by Protein A/G column (Pharmacia), and applied to affinity columns to which the appropriate synthetic peptides had been covalently linked (Pharmacia). The columns were washed with several volumes of buffered saline and then eluted with Elution buffer (pH 2.8, Pharmacia). The eluted fractions were collected into tubes containing sufficient 1 M Tris-HCl, pH 8.0, to bring their pH to 7.2. Affinity-purified antibodies and the antibody fractions that failed to bind to the affinity column were dialyzed against buffered saline containing 10mM NaN₃ and stored at 4°C.

The specificity of the antibodies was assessed by enzyme-linked immunosorbent assays (ELISA), following the method of Engvall (Methods Enzymol. 70:419-439, 1980), and by immunoblotting of the synthetic peptides separated by SDS-PAGE. The results from ELISA confirmed the specificity of the antibodies for their corresponding antigens, as did the immunoblotting.

EXAMPLE 3

In accordance with the present invention, genes such as the ones discussed above which are involved in growth and differentiation of hepatocytes will also be involved in liver repair. This is important because cirrhosis or end stage liver disease is: (1) the fourth most common cause of death in the U.S; (2) related to fibrosis and nodular hyperplasia; (3) an important risk factor for hepatocellular carcinoma; and (4) currently has no suitable medical treatment.

One such mode of treatment will be the use of the elf

protein, such as the three isoforms elf 1-3 as set forth above. In Figure 13, the ELF spectrin membrane skeleton is shown. Spectrins are rod shaped, alpha and beta subunits. Helix linked by short actin filaments at junctional complexes that include AE2, protein 4.1, myosin. This membrane skeleton attaches to the plasma membrane at 2 sites, by ankyrin and the beta subunit of spectrin.

In Figure 14, a comparison of ELF3 and 1 to Beta general spectrin is shown. Spectrins have three domains: Domain I binds Actin; Domain II binds to ankyrin; and Domain III dimerizes spectrin and gives tissue specificity. Sharp differences occur at the amino and C terminal ends. ELF1 differs from ELF3 in that it does not have an ankyrin binding domain.

Further studies have shown that the functional role of ELF may be associated with the ankyrin binding domain. Through antisense oligonucleotides to the ankyrin binding domain, ANK1 inhibits membrane associated ELF3. In addition, the inhibition of the ANK binding domain of ELF3 may result in the loss of intrahepatic bile ducts. This has a phenotype similar to Primary Biliary Cirrhosis (PBC): a disease of unknown etiology resulting in the destruction of Small and Medium sized intrahepatic Bile Ducts.

ELF has a distinct pattern of ELF expression in primary biliary cirrhosis, such as shown in Figure 15. In panel A - a decrease in membrane labeling of ELF in early PBC is shown. In panels B and D - moderately advanced PBC is shown along with an absence of membrane labeling with a concomitant increase in cytoplasmic labeling of ELF. Panel C shows the absence in ELF labeling in fibrotic cirrhotic PBC tissue, lacking hepatocytes.

Accordingly, the evidence shows that ELF1 and 3 are Beta Spectrins expressed in bile duct epithelial cells and hepatocytes. In addition, ELF3 (membrane spectrin) inhibition leads to loss of intrahepatic bile ducts in explant cultures, with increased presence of lymphocytes. Decreased membrane

labeling of ELF and marked increase in cytoplasmic labeling characterize PBC.

Further, antisense studies appear to show a primary role for ELF in PBC. Antibody studies in PBC support the role of ELF in its pathogenesis. Abnormal distribution of ELF may affect intracellular trafficking and may explain changes in AE2 expression and the resulting cholestasis seen in PBC.

Since the above reflects the mechanism by which ELF disruption results in primary biliary cirrhosis, investigations were made as to whether there were other pivotal protein interactions involved. Recent studies in smad2/3 mutants suggested the involvement of the TGF Beta pathway in PBC and the ultimate use of the elf proteins of the present invention in treating or preventing PBC and other liver diseases.

Transforming growth factor- β (TGF- β) is the major cytokine involved in organ fibrosis. It inhibits growth of hepatocytes and some hepatocellular carcinomas (HCC). The SMAD proteins serve as intracellular mediators of TGF- β and activins. TGF-Beta receptor activation involves phosphorylation of SMAD2 and SMAD3 and heteromeric complexes with SMAD4, such as shown in Figure 16. Complexes translocate to the nucleus to control expression of target genes.

Studies that shown that animals lacking smad2 die before 8.5 days of development (E8.5), and thus smad2 is required for gastrulation and mesoderm induction. Animals lacking smad3 are viable but suffer mucosal immunodeficiency.

To investigate potential genetic cooperativity in the smad gene family, we intercrossed strains lacking smad2 and smad3. These studies showed that (1) Mice doubly heterozygous for disruptions of the smad2 and smad3 genes display novel phenotypes not present in either single heterozygous; (2) Mutants die at E14 with marked liver hypoplasia; (3) 1-2% of smad2+/- smad3+/- animals survived to weaning; and (4) Mutant livers occur in 30-40% of the wild type.

Immunohistochemical analysis was conducted of

hypoplastic liver sections using antibody to alpha-fetoprotein, and the results are shown in Figure 17. In Fig. 17, Panel A shows wild type liver sections, cords of well organized and differentiated embryonic hepatocytes. Early primitive bile ducts are seen. On the other hand, Panel B of Fig. 17 shows smad2/3 mutant embryos, where hepatocytes are present, but normal cell architecture is lost. In addition, there is an absence of cell plates, and only primitive bile ducts.

Explant embryonic liver from mutants were cultured with HGF in an attempt to rescue the mutant embryos. When mutant embryonic livers were cultured in the presence of cardiac mesoderm, severe apoptosis occurred as seen in Figure 18b, compared to wild type cultures in Figures 18a and 19a with good hepatic growth.

Mutants cultured in the presence of HGF (5/50 ng/mL) show a rescue of the hepatic phenotype, with the formation of well differentiated hepatocytes, as well as primitive bile ducts as seen in Figure 19c.

These tests also showed that smad2/smad3 mutant embryos die at day 14 with profound anemia and liver hypoplasia, and hepatic stem cell proliferation in the smad mutants is dramatically reduced. Hepatocyte and erythrocyte differentiation in the smad mutants is reduced by 10%, and HGF can rescue the hepatic phenotype in explant cultures.

Accordingly, it appears that Smad 2 and 3 are essential for hematopoiesis and growth of developing liver. Gut and hepatic lineage are not altered in smad mutant mice. Smad 2 and 3 thus appear to be required for hepatic stem cell proliferation and cytoskeletal organization.

Moreover, ELF inhibition results in a phenotype with features suggestive of primary biliary cirrhosis (PBC). ELF expression is identical in smad2/3 mutants and PBC. It thus appears that smad2/3 is involved in the pathogenesis of PBC. Evidence for this includes the suppression of smad3 in PBC; Smad2 nuclear localization absent in PBC; Smad2 and 3 bind to

ELF in PBC tissue; Smad2/3 mutants have a severe but similar phenotype seen in primary biliary cirrhosis; and cytoskeletal protein interactions with ELF spectrins play a pivotal role in the pathogenesis of primary biliary cirrhosis.

It thus appears that genes such as the *elf* 1,3 proteins of the present invention and smads 2,3 which are involved in growth and differentiation of hepatocytes will be involved in liver repair seen in diseases such as primary biliary cirrhosis.

In summary, transforming growth factor- β (TGF- β) is a major cytokine involved in multiple cellular processes including differentiation, proliferation, migration, fibrosis and apoptosis, and SMAD proteins serve as intracellular signaling molecules of TGF- β and activins. In novel phenotypes with *smad2/3* intercrosses, almost all mutants died at E14 with marked liver hypoplasia and loss of primitive bile ducts. The *smad2/3* mutants were notable for a marked fall in the expression of *elf3*, a β -spectrin in accordance with the present invention. Antisense studies to *elf3* and studies in tissue from patients with primary biliary cirrhosis suggest that a crucial role for ELF in this disorder and intrahepatic bile duct formation. In addition, *smad2/3* mutants have a severe but similar phenotype seen in primary biliary cirrhosis. SMAD2, and SMAD3 bind to ELF3. These results taken together suggest that *elf* gene is an important player in intrahepatic bile duct formation and cirrhosis, and this process involves Smad2 and Smad3.

In short, ELF interactions with TGF Beta signaling molecules Smad2 and 3 are crucial for bile duct formation, as well as cirrhotic conditions such as PBC.

EXAMPLE 4

In our search for genes that are involved in liver repair, we utilized a specific cloning strategy of subtractive hybridization with embryonic liver cDNA libraries, and

identified two novel Beta Spectrins termed *elf*(4), such as disclosed above. Antisense studies utilizing cultured liver explants show a vital role of *elf*3 in hepatocyte differentiation and intrahepatic bile duct formation, and to be a marker in cirrhosis. A similar loss of intrahepatic bile ducts is noted in *smad*2/3 knockouts. Notably ELF expression is diminished in these knockouts. Furthermore, we have shown that ELF binds directly and specifically to TGF Beta signaling molecules Smad2 and Smad3. Together, these results suggest a model in which ELF interactions with Smad2 and Smad3 are pivotal for bile duct formation.

These results suggest that ELF interactions with TGF Beta signaling molecules Smad2 and 3 are crucial for bile duct formation, as well as cirrhotic conditions such as PBC. Experiments performed on these molecules have given initial information of the mechanisms associated with cirrhosis and repair and how ELF and the TGF Beta signaling pathway activate these regulatory mechanisms.

These results demonstrate that *elf* is expressed as four transcripts in the liver, a 9.0 Kb primary transcript and three secondary transcripts (5 Kb, 4.0 Kb and 2.4 Kb). In addition, studies have shown an interaction between endogenous ELF and Smad2. For this, ELF was immunoprecipitated from liver lysates and HepG2 cells using an affinity-purified anti-ELF polyclonal antibody, and Smad2 and 3 were visualized by immunoblotting with anti-Smad2 and anti-Smad3 antibody. In immunoprecipitates prepared with preimmune antisera, no Smad2 was detectable. However, in the anti-ELF immunoprecipitates, we could clearly detect Smad2 and 3 coprecipitating with ELF.

Together, these results demonstrate that ELF is a specific partner for receptor-regulated Smads 2 and 3 in the TGF β /activin signaling pathway. Our biochemical analyses of ELF/Smad2 and Smad3 interactions suggest that ELF functions upstream in the pathway and might control the subcellular localization of Smad2 and Smad3.

Other important information of relevance to the

usefulness of the elf proteins of the present invention relates to the following compounds or factors:

Transforming growth factor-Beta (TGF-Beta)

Transforming growth factor-Beta (TGF-Beta) represents an extensive family of growth and differentiation factors including activin/inhibins and bone morphogenetic proteins (BMPs) (Heldin et al., 1997), that mobilize a complex signaling network to control cell fate by regulating proliferation, differentiation, motility, adhesion, and apoptosis. TGF-Beta promotes the assembly of a cell surface receptor complex composed of type I (TbRI) and type II (TbRII) receptor serine/threonine kinases. In response to TGF-Beta binding, TbRII recruits and activates TbRI through phosphorylation of the regulatory GS-domain. Activated TbRI then initiates cytoplasmic signaling pathways to produce cellular responses. SMAD proteins together comprise a unique signaling pathway with key roles in signal transduction by TGF-Beta and related factors. The founding member of the SMAD family, Mothers against dpp (Mad) was identified as a dominant enhancer of weakly mutant alleles of decapentaplegic, a BMP homologue in *Drosophila melanogaster* (Raftery et al., 1995; Sekelsky et al., 1995). Genetic screens in *Caenorhabditis elegans* for mutant phenotypes like those observed for Ser/Thr kinase receptors daf-1 and daf-4 revealed three genes, sma-2, sma-3 and sma-4, with homology to Mad (Savage et al., 1996). At present, nine vertebrate SMADs have been identified (Attisano and Wrana, 1998). They are characterized by homologous regions at their N- and C-termini known as Mad homology (MH)-1 and MH-2 domains, respectively.

Three classes of Smads with distinct functions have been defined: the receptor-regulated Smads, which include Smad1, 2, 3, 5, and 8; the common mediator Smad, Smad4; and the antagonistic Smads, which include Smad6 and 7 (Heldin et al., 1997; Attisano and Wrana, 1998; Kretzschmar and Massague, 1998). Receptor-regulated Smads (R-Smads) act as direct substrates of specific type I receptors, and the proteins are

phosphorylated on the last two serines at the carboxyl terminus within a highly conserved SSXS motif (Macias-Silva et al., 1996; Abdollah et al., 1997; Kretzschmar et al., 1997; Liu et al., 1997b; Souchelnytskyi et al., 1997). Regulation of R-Smads by the receptor kinase provides an important level of specificity in this system. Thus, Smad2 and Smad3 are substrates of TGF β or activin receptors and mediate signaling by these ligands (Macias-Silva et al., 1996; Liu et al., 1997b; Nakao et al., 1997), whereas Smad1, 5 and 8 are targets of BMP receptors and propagate BMP signals (Hoodless et al., 1996; Chen et al., 1997b; Kretzschmar et al., 1997; Nishimura et al., 1998). Once phosphorylated, R-Smads associate with the common Smad, Smad4 (Lagna et al., 1996; Zhang et al., 1997), and mediate nuclear translocation of the heteromeric complex. In the nucleus, Smad complexes then activate specific genes through cooperative interactions with DNA and other DNA-binding proteins such as FAST1, FAST2, and Fos/Jun (Chen et al., 1996, 1997a; Labbe et al., 1998; Zhang et al., 1998; Zhou et al., 1998). In contrast to R-Smads and Smad4, the antagonistic Smads, Smad6 and 7, appear to function by blocking ligand-dependent signaling (reviewed in Heldin et al., 1997).

Genetic analysis in *Drosophila melanogaster* and *Caenorhabditis elegans*, as well as T β RII and SMAD mutations in human tumors, emphasizes their importance in TGF-Beta signaling. Collectively, these factors constitute a communication network exploited by TGF-Beta family members to regulate gene expression, and suggest a paradigm in which signaling pathways activated by ligand binding and operating in parallel, converge at target promoters to produce ligand specific responses.

Receptor Interacting Proteins

Proximal signaling events coupling TGF-Beta receptor activation to biological responses involves proteins, such as FKBP12, *Drosophila* inhibitor of apoptosis (DIAP)-1 and -2

(Oeda et al., 1998), and T β RI associated protein (TRAP)-1 and -2 (Chang et al., 1998). The WD-domain protein TRIP-1 and T β RII (Chen et al., 1995), that directly bind the receptor complex, FKBP12, a binding protein for the macrolide immunosuppressant FK506, interacts with a Leu-Pro motif in the GS-domain of T β RI and other type I receptors (Wang et al., 1996, Chen et al., 1997). (Wang et al. 1996).

Phosphorylation of R-Smads by the type I receptor is essential for activating the TGF β signaling pathway (Heldin et al., 1997; Attisano and Wranga, 1998; Kretzschmar and Massague, 1998). However, little is known of how Smad interaction with receptors is controlled. Recently, a novel Smad2/Smad3 interacting protein that contains a double zinc finger, or FYVE domain, and has been identified called SARA. SARA recruits Smad2 into distinct subcellular domains and that SARA colocalizes and interacts with TGF β receptors. TGF β signaling induces dissociation of Smad 2 from SARA with concomitant formation of Smad2/Smad4 complexes and nuclear translocation.

Moreover, deletion of the FYVE domain in SARA causes mislocalization of Smad2 and inhibits TGF β -dependent transcriptional responses. Thus, SARA defines a component of TGF β signaling that functions to recruit Smad2 to the receptor by controlling the subcellular localization of Smad.

SMAD Domain Functions and Regulation by Intrinsic and Extrinsic Mechanisms

The MH2 region of pathway-restricted SMADs is involved in protein-protein interactions, particularly with other transactivating factors, such as interaction between Smad2 and the winged-helix transcription factor FAST-1 (Chen et al., 1996). Similarly, between Smad3 and the transcriptional coactivator CBP/p300 (Fen et al., 1998; Janknecht et al., 1998). Additionally, the MH2 domains are responsible for homomeric and heteromeric interactions between SMADs (Zhang et al., 1997).

SMAD proteins reside in the cytoplasm, and upon stimulation translocate to the nucleus as part of an oligomeric complex (Attisano and Wrana, 1998). The observation that MH1 domain deletion from Smad2 results in constitutive nuclear localization (Baker and Harland, 1996) suggests an intrinsic inhibitory role for the MH1 domain in signaling by pathway-restricted SMADs. Intrinsic inhibition of SMAD function mediated by the MH1 domain is relieved by agonist induced phosphorylation of the -SSXS motif, which presumably antagonizes the intramolecular MH1-MH2 interaction (Eppert et al., 1996; Schutte et al., 1996).

The MH1 domain is also involved in direct DNA binding. The MH1 domain Mad is necessary and sufficient for binding to the "quadrant" enhancer of the vestigial (vg) wing patterning gene in *D. melanogaster* (Kim et al., 1997). Similarly, Mad binds to the Dpp response element in the Ultrabithorax (ubx) promoter via its MH1 domain. Direct interaction of Smad3 and Smad4 with a CAGA-box, a DNA element repeated three times in the TGF-Beta responsive regions of the plasminogen activator inhibitor (PAI)-1 promoter has also been shown. This interaction requires the MH1 domain of both Smads. Smad3 additionally requires either agonist stimulation or MH2 domain deletion. In a basal state, MH1 and MH2 domains provide intrinsic, reciprocal inhibition that is liberated by receptor activation and -SSXS phosphorylation.

Involvement of extrinsic regulatory pathways such as the ERK MAP kinase pathway also contribute to SMAD regulation. In response to mitogenic growth factors such as epidermal growth factor (EGF) that signal through receptor tyrosine kinases (Denhardt, 1996), ERK-mediated phosphorylation of target transcriptional regulators contributes to the mitogenic influence of these factors. Recently, multiple serine residues in the linker region of Smad1 were shown to be phosphorylated by ERK, both in vitro and in vivo in response to EGF (Kretzschmar et al., 1997b). While phosphorylation of Smad1 by ERK was independent of -SSXS phosphorylation and did

not effect association with Smad4, it did antagonize nuclear translocation of the SMAD oligomeric complex in response to BMP stimulation. ERK mitogenesis may involve simultaneous potentiation of growth promoting pathways and attenuation of growth inhibitory pathways.

In addition, Ca²⁺ dependent interaction between calmodulin (CaM) and several SMAD family members has been described (Zimmerman et al., 1998). CaM bound the N-terminal half of Smad2 between residues 76 and 208. Both CaM coexpression and a CaM-specific antagonist suggested a negative regulatory role for CaM in both activin and TGF-Beta signaling in a transient assay. Given the wide array of factors regulated by CaM either directly via protein-protein interaction or indirectly by CaM-dependent kinases, it is attractive to speculate that CaM influences SMAD protein function in response to agents that regulate intracellular Ca²⁺ flux.

Finally, other groups of proteins in which mammalian homologues may be important in liver formation due to interactions with the ELF proteins of the present invention are the family of growth factors produced by the fat body and which are active on *Drosophila* imaginal disc cells, such as disclosed in Kawamura et al., Development 126:211-219 (1999), incorporated herein by reference.